Journal of Buddhist Ethics
ISSN 1076-9005
http://blogs.dickinson.edu/buddhistethics/
Volume 21, 2014

Recent Buddhist Theories of Free Will: Compatibilism, Incompatibilism, and Beyond

Riccardo Repetti

Kingsborough College, CUNY

Recent Buddhist Theories of Free Will: Compatibilism, Incompatibilism, and Beyond

Riccardo Repetti [1]

Abstract

This is the fourth article in a four-article series that examines Buddhist responses to the Western philosophical problem of whether free will is compatible with "determinism," the scientific doctrine of universal lawful causation. The first article focused on "early period" scholarship from the 1970's, which was primarily compatibilist, that is, of the view that the Buddhist conception of causation is compatible with free will. The second and third articles examined "middle period" incompatibilist and semi-compatibilist scholarship in the remainder of the twentieth century and first part of the twenty-first. The present article examines work published in the past few years. It largely agrees that Buddhism tacitly accepts free will (al-

[1] Department of History, Philosophy, and Political Science, Kingsborough College, CUNY. Email: rrepetti@kingsborough.edu.

> though it also explores an ultimate perspective from which the issue appears moot), but mostly divides along compatibilist and incompatibilist lines, mirroring Theravāda and Mahāyāna Buddhist perspectives, respectively. Of the writers I emphasize, Gier and Kjellberg articulate both perspectives; Federman and Harvey advocate Theravāda compatibilism; and Wallace argues that although determinism and free will are incompatible, subtle complexities of Mahāyāna Buddhist metaphysics circumvent the free will and determinism dichotomy. Although the present article focuses on these writers, as the culminating article in the series it also draws on and summarizes the other articles in the series, and directs the reader to other recent period works that, due to space constraints, cannot be reviewed here.[2]

Overview and Disclaimer

There are few passages, if any, in canonical or authoritative Buddhist texts that explicitly address anything remotely resembling the philosophical problem of free will. This series of four articles is intended as an in-depth guide to the secondary literature on Buddhism and free will, a fairly small and relatively recent body of writings that began roughly within the past half century. Although these four articles appear in this

[2] I would like to thank Daniel Cozort, Peter Harvey, and an anonymous reviewer for their many helpful questions, clarifications, objections, and editorial suggestions, and Asaf Federman for his review of the section on his work. The writing of this article was supported in part by my participation in the National Endowment for the Humanities 2012 Summer Institute, "Investigating Consciousness: Buddhist and Contemporary Philosophical Perspectives," in Charleston, South Carolina.

journal, the intended audience is not only scholars and students of Buddhist ethics, but also Western philosophers who for the most part have only recently begun to take Buddhism seriously and to try to mine it for ideas, arguments, theories, and alternative ways of knowing.

This dual audience complicates matters. On the one hand, many Buddhist scholars are not only unfamiliar with the Western philosophical problem of free will and its many iterations in the hands of soft versus hard determinists, libertarians, compatibilists, incompatibilists, and semi-compatibilists, but are also unfamiliar with what other Buddhist scholars have written about free will. (I will explain these terms below.) Even much of the work of Buddhist ethicists betrays varying degrees of unfamiliarity with relevant doctrinal distinctions pertaining to divergent forms of Buddhism as well as to divergent theoretical perspectives within Western philosophy. On the other hand, Western philosophers who wish to explore what Buddhists think about free will might be significantly unfamiliar with basic Buddhist ideas such as Dharma, dependent origination, and the twelve-linked chain, or classic examples such as the chariot and its parts, not to mention doctrinal differences within the various traditions and schools of Buddhist thought. (I will explain these ideas briefly below.)

Some of the work I will review in this article, for example, by Gier and Kjellberg, seems to presuppose an extensive background in Buddhist thought. This is despite the fact that it is contained in an edited collection of otherwise entirely Western analytic philosophical articles on free will, presumably targeting a Western audience. I wrote this series of articles in order to aid scholars on both sides of the Western/Buddhist divide in their attempts to understand the many often confusing if not conflicting claims to be found within the extant secondary literature.

Review of Early- and Middle-period Scholarship

My analysis of early-period and middle-period scholarship, in the earlier articles in this series,[3] supports the idea that the Buddha's teaching is mental-freedom-oriented and rests upon his doctrine of *pratītya-samutpāda* (dependent origination), the general doctrine that every event originates in dependence on multiple simultaneous and/or previous conditions, and the particular application of this generalization to the *nidānas* (twelve causal links in the chain of suffering). As the Buddha stated, "When that is present, this comes to be; on the arising of that, this arises. When that is absent, this does not come to be; on the cessation of that, this ceases" (MN I 262 ff.; SN II 28). Note that, on its face, this description does not assert that causation is anything more substantive than a conditional relationship: when this condition is satisfied, that condition occurs.

For ease of reference, let's call this minimalist description the Buddha's "conditionality formula." Determinism presupposes that a metaphysically maximalist nomological (lawful) necessity—what Jay Garfield describes as the "cement of the universe"[4]—serves as the truth condition for any instantiations of the conditionality formula. Thus, few Buddhists would be willing to equate minimalistic conditionality with maximalistic determinism; Charles Goodman appears to be one noteworthy exception.[5] By way of objection, however, it does not follow that just because instantiations of the minimalist conditionality formula are true that there is no maximalist nomological necessity that accounts for their truth. Reluctance to acknowledge something does not evidence that it

[3] Repetti ("Earlier," "Reductionism," and "Hard Determinism"); see, also, Repetti ("Meditation").

[4] Garfield ("Causality").

[5] Goodman ("Resentment"); for a critical review, see Repetti ("Hard Determinism").

does not exist. What makes it reliably the case that the conditionality formula is true? That it makes sense to ask this question suggests that something stronger than mere coincidence must be in play. Although it is conceivable that everything that appears to involve a substantive causal relationship might amount to no more than a "constant conjunction" between event pairs, as David Hume famously suggested (and thought was all that empirical observation served up, technically),[6] it is implausible in the face of the widespread contemporary scientific understanding of causal laws as exceptionless nomological generalizations.[7] If nomological necessity does ground conditionality, Buddhist conditionality would ultimately reduce to a partly opaque form of determinism.

---

[6] Hume (*Treatise*).

[7] On the understanding of causation that has the broadest consensus in the Western philosophical and scientific communities (Woodward and Hitchcock, "Generalizations"), laws support an indefinite number of counterfactuals, that is, statements of the form, "had x happened, then y would have happened," and/or subjunctives, statements of the form "were x to happen, then y would happen," both of which imply something stronger than mere Humean constant conjunction insofar as they imply that all past cases that would have satisfied the antecedents of these conditionals would also have satisfied the consequents of these conditionals, that all future cases that satisfy the antecedents of these conditionals will also satisfy the consequents of these conditionals, and that all possible cases that hypothetically satisfy the antecedents of these conditionals also satisfy the consequents of these conditionals. For example, let an instantiation of the subjunctive "were x to happen, then y would happen" be "were a particular grain of salt to be submerged in water, then it would dissolve." This subjunctive conditional is true, and it remains true even if that particular grain of salt is never submerged in water. The scientific understanding of the nomological necessity involved in these counterfactual and/or subjunctive conditionals attributes the solubility of salt to its crystalline atomic structure and the structure of water molecules, in which case the scientific understanding of the causation of salt's solubility and/or its actual dissolving in water requires an interpretation of causation that is not only stronger than Humean constant conjunction, but also stronger than Buddhist conditionality, despite the fact that the counterfactuals and subjunctives involved in the explication of laws are conditionals themselves.

Early-period scholars clearly resisted equating Buddhist causation (dependent origination) with either a "rigid" determinism or a "chaotic" indeterminism, opting for a "middle way" between both.[8] Some middle-period scholars, however, accepted an outright "hard determinism,"[9] the view that determinism is true and entails an invariable series of events, entailing that whatever happens is inevitable and that we lack free will.[10] Others opted for what is nearly the same view, a "semi-compatibilism" according to which determinism remains "hard" in terms of ruling out agential autonomy at the level of *ultimate* reality but is "soft" in terms of being compatible with the sort of free will (in the

[8] Repetti ("Earlier").

[9] Repetti ("Hard Determinism").

[10] Hard determinism resembles fatalism, but differs from it in one crucial respect. What these doctrines share is that both entail a singular future series of necessary events, no member-event of which is agent-alterable or agent-avoidable. The crucial difference is that determinism explains this necessity *causally*, such that event E_a causes event E_b, E_b causes E_c, and so on, whereas fatalism simply asserts the *acausal* necessity of all events E_a, E_b, E_c, . . . , E_n. Hard determinism seems initially, therefore, to be a more rational doctrine. Ironically, however, some fatalists at least entertain the possibility that agents might actually be able to make (unfortunately futile) different choices and efforts to alter what is fated, whereas hard determinists deny that agents are even able to engage in any deliberations, or make any choices or efforts, other than those that are pre-determined. Some hard determinists concede that determinism alone does not rule out the truth of counterfactual hypotheticals about how an agent could have done otherwise had motivational and related conditions been otherwise, but they insist that because agents cannot bring it about that pre-agential conditions are ever otherwise, agents lack any ability to alter the pre-determined future. That is, they think this "ability" is impotent, null. To my understanding, this is a key point of disagreement that differentiates hard from "soft" determinism, which latter view embraces the idea that certain forms of agential control, although determined, render undesirable outcomes sufficiently evitable for purposes of moral responsibility.

moral-responsibility-entailing sense) that we attribute to persons at the level of *conventional* reality.[11]

My own view is that, if we assume that dependent origination is deterministic, then the Buddha's teaching should be characterized as "soft determinism," the view that determinism is true, that free will (in the moral-responsibility-entailing sense) is real, and that determinism does not entail inevitability (because knowledge of volitional causes and effects renders undesirable actions and outcomes avoidable). The Buddha rejected the idea that actions are inevitable, and he rejected fatalism on that ground (DN I 249-253).[12]

Arguably, the Buddha seems to have tacitly accepted a minimalistic sense of "free will" involving wholesome volitional regulation in connection with progress along the Buddhist path. Because mindful volitional regulation fosters mental freedom and mindless unregulated volition fosters mental bondage, I think the Buddha would consider the only *relevant* type of free will one that involves volitional regulation, as

---

[11] Repetti ("Reductionism"). The meaning of the Buddhist distinction between "ultimate" and "conventional" reality is disputed among most schools of Buddhist thought (Thakchoe, "Truths"), and thus is itself worthy of at least a book-length treatment. A sketch will suffice for now: The distinction maps loosely on the appearance/reality distinction, such that conventional reality is how things appear to the unenlightened mind, namely, they appear to be substantive and independently existing, whereas ultimate reality is how things are independently of conceptualization altogether, namely, they are insubstantial and exist only dependently. The meaning of this distinction will come into greater focus below.

[12] *Dīgha Nikāya* (DN), "Long Discourses." Goodman ("Resentment") rightfully disagrees with the equation of hard determinism and fatalism, because fatalism is *acausal*, as I noted earlier. However, Harvey ("Freedom") and Federman ("Buddha") each argue forcefully that hard determinism and fatalism share the implication that actions are inevitable, which I also noted earlier, that the Buddha rejected the idea that actions are inevitable, and thus that the Buddha would have rejected hard determinism.

opposed to the sort of unregulated volitional expression associated with free will in the West—the freedom to do as one pleases or to express one's volitions spontaneously in actions. Regulated volitional expression may be described as "controlled" will and unregulated volitional expression may be described as "unrestrained" or "free" will, where "free" implies "spontaneous." According to Buddhism, it is regulated will that leads to the central Buddhist goal of liberation or mental freedom—freedom precisely from the sort of mental bondage caused by unregulated will.[13]

There is a related compatibilist sense of "free" will that consists in the will's being proximately controlled by the agent (such as when choices and actions are voluntary, as opposed to involuntary), and a semi-compatibilist sense of "free" will that consists in the will's being reason-responsive (such as when the agent has a moral reason to refrain from a certain action and does so). But reason-responsiveness grounds the designation "free," not based on considerations about whether the will is regulated or unregulated, but on the Kantian idea that moral responsibility implies free will. That is, it is grounded on the idea that "'ought' implies 'can'": if reason-responsiveness is sufficient for moral responsibility, then there is a minimal sense in which the reason-responsive agent must be able to behave morally, even if the fact that the agent is reason-responsive is determined and thus the agent cannot be truly autonomous in the sense that full-blooded autonomy requires that the agent's choices not be determined. Proximate agential control over the will and reason-responsiveness are, according to compatibilists and semi-compatibilists, respectively, sufficient for purposes of moral responsibility.

---

[13] This line of thought calls to mind an interesting parallel: between unregulated volitional expression as a chief cause of individual suffering and unregulated consumerism as a chief cause of collective suffering, but that is a topic for another inquiry.

Again, the Buddha would likely only consider the type of free will that is relevant to the Buddhist path—and thus the only form of free will that is worth cultivating—that of regulated will, on the one hand, and he would likely consider that type of free will that consists of unregulated will to be antithetical to the Buddhist path, on the other hand. For this reason, in arguing for what I take to be the most reasonable form of free will that is implicitly relevant to Buddhism, I restrict my focus to the relationship between regulated will and mental freedom, because it is only by reference to liberation that any form of free will—even regulated will—becomes relevant to Buddhism at all. Thus, when I refer to the Buddhist notion of free will, I mean the regulated will that is cultivated on the path to liberation.

It must be emphasized that this interpretation and emphasis does not necessarily involve the same conception of free will in play in various Western or Buddhist accounts of the subject, not to suggest that there is a single conception of free will in play in both or either. Rather, it is simply the interpretation that I am arguing for, perhaps on slightly revisionist grounds. I think that both the compatibilist and semi-compatibilist senses of free will—for example, the will's being proximately controlled by the agent and its being reason-responsive, respectively—may be accounted for in Buddhist terms along the lines of volitional regulation. Most of us, unfortunately but realistically, possess very little of this sort of free will: control over our own volitions. Buddhist practitioners possess it increasingly, however, covariant with their cultivation of mental freedom, and typically through meditative and related cognitive disciplines.[14]

My review of early- and middle-period scholarship supports the following *remedial* distinctions, then. Buddhism differentiates between

[14] See Repetti ("Meditation").

the agency of *āryas* (noble persons), that is, Buddhists with some degree of enlightenment (those at the rank of stream-enterers, once-returners, non-returners, and *Arhats*,[15] and/or those well established in the strong insight that leads to each of these), on the one hand, and non-*āryas*, on the other. For ease of reference, I will refer loosely to the latter group as "worldlings."[16] Although there are exceptions, worldlings typically possess varying degrees of relatively minimal or merely potential agency, in light of their habitually unrestricted volitional expression, which strengthens mental-bondage-fostering dispositions; Buddhist practitioners, however, cultivate volitional regulation, which increases mental freedom. Because volitional regulation is consistent with determinism, it does not entail "libertarianism," the view that we have free will and that, because free will and determinism are incompatible, determinism is false.

Fully enlightened persons enjoy full mental freedom, and even though they transcend ego-volitional behavior, they possess some sort of

---

[15] Stream-enterers have attained the first of the four stages of enlightenment in significantly recognizing the truth of the Dharma and committing to the Buddhist path, once-returners have attained the second stage in reaching an initial state of spiritual transformation that entails they will reincarnate once more in human form at most before attaining enlightenment, non-returners have attained the third stage in reaching an advanced state of spiritual transformation that entails that they will not reincarnate again in human form before attaining enlightenment, and *Arhats* have attained the fourth, final stage of full enlightenment. Christopher Gowans ("*Philosophy*," 13) coined the neologism, "stream-observer," to identify another relevant category: one who is examining the Dharma to see if Buddhism's main claims and soteriological strategies are credible and worthy of pursuit. For the many Western practitioners of Buddhist meditation who seem to be traveling significantly but not fully along the Buddhist path (as an informal experiment that tests the teachings through direct experience), these categories may be fuzzy or overlap.

[16] "Worldlings" is restricted to anyone that has not attained the first stage of realization, whether or not they are Buddhist practitioners.

self-regulative control in that they are masters of volitional regulation, arguably a form of free will.[17] If they lack ego-volitions, arguably, *they* perform no actions attributable *to them* and therefore *they* are not determined. Here, the word "they" is ambiguous: one sense involves reference to an ontological agent that doesn't exist or function in actions and the other to a psychological agent that does exist and function in actions. Thus, in the psychological sense, "they" do perform actions, technically, such as advising, teaching, walking, and so on, but their actions are *impersonal* manifestations of compassion, untainted by ego-based

[17] This claim may seem puzzling on its face because if there is no longer any illusion about a genuinely existing ontological self or agent, the idea of self-regulation seems impossible. But what is puzzling about it, arguably, might just be our own association with an ontological self: surely an enlightened being can have a highly functional psychological self—effective agency—with no illusions about its ontological status. Buddhist lore is full of such highly effective *Arhats* and bodhisattvas. Perhaps the intuition is that absent any ego-based (delusional) volition the *Arhat* is not the causal nexus of her actions, but rather the universe is, not unlike the way spontaneous action is conceived in Daoism: control or regulation, on this view, is only necessary or even possible when there is a perceived conflict between volitions and/or preferences, but in *Arhats* there cannot be any such conflict. But this seems to overly idealize enlightenment in a way that ignores real possibilities, such as the one illustrated by the Buddhist story of the *Arhat* who met a close friend of his deceased father's one day who insisted on the monk's drinking with him, to excess (Rinpoche, *Chariots*, 22), to use just one example. The idea is that even *Arhats* can be presented with compelling volitions, their own or others', and despite how typically effortless or spontaneous it may be to navigate them, they still need navigating, and sometimes that navigation might have to be nonspontaneous or deliberative. Conceivably, an *Arhat* could be drugged, and presented with elaborately compelling pressures that require volitional regulation; or a neuroscientist could kidnap the *Arhat* and insert a remotely-controlled chip in its brain that generates powerful unwholesome volitions. The point is that an *Arhat* is arguably capable of exercising volitional regulation, should the need ever arise, even if the need typically never arises. By analogy, salt remains soluble even if it will never be submerged in any liquid.

volition or any ontological-self illusion, and thus they generate no karma.

If we assume that dependent origination is sufficiently deterministic, we can call this model of impersonal enlightened action "*virtually-libertarian*-soft-determinism" because in the absence of any ontological self there can be nothing to undermine maximally effective psychological agency, the functional equivalent of a psychological self's autonomy in the absence of belief in any ontological self. Enlightened action provides an analogue of the limit case of autonomy that the libertarian thinks obtains and which implies that determinism is false. But this model is consistent with determinism, hence "soft," and thus only *virtually* libertarian. Buddhist practitioners are typically *relatively-soft*-determined—by degree, relative to their degree of regulated volitional development. And worldlings are typically *virtually-hard*-determined—by degree, relative to their largely unregulated volitional underdevelopment and almost utter lack of volitional control—but *capable* of being *relatively*-soft-determined upon hearing the Dharma.[18]

Some key issues remain open following my review of early- and middle-period scholarship. One is the interpretation and impact of the *anātman* (non-self) theory, the Buddhist doctrine about the metaphysical/ontological insubstantiality of the self. Until fairly recently, this was a uniquely Buddhist view: what people take to be a substantial and permanent self-essence to a person actually is a mereological (part/whole) fabrication of the aggregates—of form (matter), sensation, perception, volition, and consciousness—that constitute us and account for all our behavior and experiences. Thus, despite the clarification afforded to the

[18] Of course, many worldlings cultivate greater degrees of volition regulation ability for non-Buddhist moral, religious, or spiritual reasons, but to the extent that they do there is likely some overlap with the bases of their behavior and the teachings of Buddhism.

issue by the distinction between the psychological and ontological senses of the self, it remains a puzzling question whether something that is a non-self can exhibit self-regulation (volitional control). Some Buddhist philosophers, such as Charles Goodman,[19] patently deny that a non-self can have any abilities at all, much less autonomous control over them. Others, such as Mark Siderits,[20] think that the "two truths" doctrine enables Buddhists to both deny that there is a substantive self and to assert that persons ordinarily exhibit free will.[21] According to this doctrine, often called upon to address philosophical puzzles in Buddhism, at the conventional reality level free will (possibly in the moral-responsibility-entailing sense, but clearly in the volitional regulation sense) matters, but at the ultimate reality level there is no substantive self and thus no (metaphysically real) "agency."

Early-period scholarship mostly shares a neutral, middle-path position between "rigid" (hard) determinism and "chaotic" (libertarian) indeterminism that I call "wiggly-determinism" and which I have argued is unsuccessful.[22] Goodman,[23] one middle-period scholar, simply embraces hard determinism, whereas Siderits,[24] another middle-period scholar, adduces a "paleo-compatibilist" position between hard determinism and libertarian indeterminism by parsing each at different levels of reality

[19] See Goodman ("Resentment"); see also a critique in Repetti ("Hard Determinism").

[20] See Siderits ("Paleo-compatibilism"); see also a critique in Repetti ("Reductionism").

[21] The "two truths" is another description for the distinction between ultimate and conventional reality. See Thakchoe ("Truths") for an excellent introduction to the complexities of this doctrine.

[22] Repetti ("Earlier").

[23] Goodman ("Resentment").

[24] Siderits ("Paleo-Compatibilism").

(where the levels are what are compatible). I have argued that Siderits's position amounts, on analysis, to a form of semi-compatibilism.[25]

Most early- and middle-period scholars sought a neutral position that might apply to all Buddhist traditions. Most recent-period scholars endorse more tradition-specific non-neutral positions. Nicholas F. Gier and Paul Kjellberg[26] show that the Theravāda tradition tends to be compatibilist whereas the Mahāyāna tradition tends to be significantly incompatibilist. Asaf Federman and Peter Harvey agree that the Theravāda tradition is compatibilist. B. Alan Wallace argues that although determinism is incompatible with free will, the anti-realism of the Mahāyāna perspective transcends the dichotomy of free will versus determinism altogether. Let us begin our review of recent-period scholarship with an analysis of the views of Gier and Kjellberg.

Gier and Kjellberg: Pāli versus Mahāyāna Buddhism[27]

Gier and Kjellberg's intellectually rich contribution to this discussion ("Buddhism") is challenging for two reasons: [28] first, it attempts to en-

[25] Repetti ("Reductionism").

[26] Gier and Kjellberg ("Buddhism").

[27] For Western readers new to scholarly Buddhism, Pāli is the language used to inscribe the original Buddhist canon. The Pāli Canon is the only fully surviving early Canon in an Indic language. Parts of other such canons survive in Chinese and Tibetan translation; later Indian Buddhist texts survive mostly in Sanskrit. Theravāda is the surviving form of Pāli Buddhism. Gier and Kjellberg use the term "Pali" (an Anglicized "Pāli") somewhat ambiguously, sometimes to refer to elements of the Pāli Canon, but other times intending what is normally denoted by the term "Theravāda." Because I quote their usage extensively, I will maintain "Pāli" for their use of the term when reviewing but not quoting their ideas.

[28] Gier and Kjellberg ("Buddhism").

gage a Western philosophical audience, and second, it shows how philosophically complex are the issues connected with free will in light of the divergence between earlier and later Buddhist doctrines. Although it may be overly ambitious, it is an important piece because it largely succeeds. Their article is situated within an edited collection of articles on the problem of free will,[29] the others of which approach the problem entirely from a Western analytic philosophical perspective, making no significant mention of Buddhism or even of Asian philosophy. Many of Gier and Kjellberg's ideas are directed toward what other contributors to that volume have said. This dimension of their thought helps explain why they address certain matters that might not seem relevant to a Buddhist audience.

Gier and Kjellberg open by paraphrasing the Buddha in a way that highlights the complexity of the Buddhist position on free will and moral responsibility: "There is free action, there is retribution, but I see no agent that passes out of one set of momentary elements into another one, except the [connection] of those elements" (277).[30] This quotation supports the idea that the Buddha does not deny the sort of voluntary action typically associated with moral responsibility, but simultaneously expresses the paradox that there could be intentional action without any diachronic (non-momentary) agent. There is a *prima facie* tension between these ideas, but Gier and Kjellberg do not resolve it. They present divergent Buddhist views without choosing among them. All I hope to achieve here is to analyze, elucidate, and pose problems for selective elements of their overview.

Engaging with one Western analytic understanding of free will that rests on the distinction between mere desires and preferences (as

[29] Campbell, O'Rourke, and Shier (*Determinism*).

[30] Quoting Stcherbatsky (*Logic*), 133.

expressed by Keith Lehrer in the same volume but traceable to Harry Frankfurt),[31] Gier and Kjellberg argue that the "Buddhist is an agent of her preferences . . . rather than being just a passive victim of her desires" (283). There is something right about one part of this claim, because through mindfulness and related disciplines Buddhists aim to cultivate detachment from—and to a certain extent control over—certain kinds of desires, some of which, from the perspective of the Buddhist path, may be parsed as preferences.[32] However, the other part of this claim seems at odds with the Buddha's denial of an agent (in their opening quote), a problem they will address later on.[33] They assert that Buddhist "determinism," unlike standard Western interpretations, is not linear: "Buddhist causality . . . is seen as a cosmic web of causal conditions rather than linear mechanical notions of push-pull causation" (283). To them, non-linearity in causality evades the fatalistic sting of linear determinism, according to which event A necessarily and inevitably causes event B, B necessarily/inevitably causes C, C causes D, and so on, leaving no wiggle room for free will.

---

[31] Lehrer ("Freedom"). Earlier on, Frankfurt ("Freedom") differentiated between base-level desires and higher-order desires, which comes close to the same thing: if I prefer desire D1 over desire D2, then I have a higher-order desire D3 for D1.

[32] See Repetti ("Meditation") for a complex development of this idea.

[33] Gier and Kjellberg make a number of incautious generalizations in their introductory section. For example, they claim that "While the issue of free will does not arise in Buddhism, it is indisputable that it embraces a universal determinism" (283), but our discussion about the minimalism of conditionality above stands in opposition to this interpretation. In support of their determinism claim, they say Buddhism asserts that "every effect, without exception, has a cause," but nobody doubts this tautology. They must have meant to say every *event* has a cause, but that claim is neither obviously true, nor universally agreed upon (for example, in quantum mechanics). And, as they will later on acknowledge, some Buddhists view causation in minimalistic terms as a form of conditionality that does not seem deterministic at all.

However, it is doubtful that non-linear causality circumvents inevitability: if everything *co-influences* everything else, it seems to make more sense to think that each multi-directional causal arrow remains linear than it does to think that things become metaphysically fuzzy, as in wiggly-determinism. On this interpretation, arguably, there are indefinitely many more linear causal arrows in Buddhist causal interdependence than there are in standard determinism. Causal interdependence simply multiplies and over-determines the inevitability of determinism—yielding a kind of *multi-directional* Buddhist mega-determinism in which every particle/event is maximally necessitated by every other one, rather than an inchoate wiggly-determinism.[34]

Gier and Kjellberg invoke the early-period Buddhists' wiggly-determinism conception, however, when they quote the Buddha to say that if there were a one-to-one karmic relationship, liberation would be impossible, adding: "They who know causation . . . know the Dharma"

[34] Using simple gravitational attraction as an illustration, compare a simple linear determinism whereby event A causes event B, on the one hand, and a complex, interdependent determinism whereby a single particle P exerts a causal—say, gravitational—force on all other particles in the universe while simultaneously all other particles in the universe exert a counter-gravitational force on particle P, and so on for every particle in the universe with respect to every other one, on the other hand. In the simple case where A causes B, there is a straightforward linear direction of influence from A to B, whereas in the interdependence of mutual gravitational influences, there are linear influences from each particle in the universe to and from each other particle in the universe. Although this is a much more complex structure of influences, within this complex structure it is still possible in principle to discern and map linear connections from and to each particle in the universe. It is in this sense that interdependence does not entail nonlinearity or nondeterminism, but rather is consistent with mega-linearity or mega-determinism. The difference, then, is not between a case of linearity versus nonlinearity, but between singular linearity (a single line of causal influence from one particle to another) and multiple linearity (indefinitely many lines of causal influence from all particles onto each particle).

(284).[35] But this interpretation seems to conflate causation with karma. Some Buddhists—some Yogācāra idealists, for example—might argue, however, that because all conditioned phenomena are mind-dependent, they are volitional and hence karmic. Apart from this minority view,[36] however, karma appears *prima facie* to be a *subcategory* of causation that the Buddha restricted to the volitional: "Volition, intention, O Bhikkhus, is what I call Kamma" (AN III 415).[37] Admittedly, however, nothing in the Buddha's quotation here is *inconsistent* with the Yogācāra view that everything conditioned is conceptualized and thus intentional, volitional, and thus karmic.

Gier and Kjellberg seem to acknowledge the more restricted interpretation when they state: "The definition of *karma* as volitional action is not only good Pali Buddhism but it is also the position of the great Mahayanist philosopher Vasubandhu: '*karma* is will (*cetanā*) and voluntary action (*cetayita karanam*).'"[38] Of course, nothing in this quotation from Vasubandhu is inconsistent with the Yogācāra view either. I cannot resolve these conflicting interpretations here, but I can provide an analysis that will circumvent the disagreement.

Thus, the Yogācāra views in question here are that all conditioned phenomena are conceptualization-dependent and that all that is conceptualized involves mental formations (*saṃskāras*: volitions). Together, these views ground the Yogācāra equation of karma and causation. Alternately put, if causal relations are conceptually embedded and

---

[35] Quoting DN I 249-253.

[36] Peter Harvey deems the question whether causation and karma are identical worthy enough to warrant significant critical discussion of the issue based on canonical sources ("Freedom").

[37] *Aṅguttara Nikāya* (AN), "Further-factored Discourses."

[38] Gier and Kjellberg (301 fn. 16), citing Stcherbatsky (*Conception*), 32.

all conceptualization involves mental formations (volition), then all conditioned or causal phenomena are karmic. From non-Yogācāra Buddhist perspectives as well as from the Western analytic philosophical perspective, not everything that is intentional or mental is volitional because, for example, simple perceptual experiences, such as one expressed by the statement "that blotch is red," are intentional because they are *about* something other than themselves (this one is about *that blotch*). But there is no phenomenological element, nor any conceptual implication, in this sort of color-perception-involving intentionality of any kind of *volitional* constituent, construed as some form of teleology, conation, desire, want, need, aim, end, goal, purpose, and so on.[39] Thus, even though everything volitional is intentional (because everything volitional aims at, and thus is *about*, some object or experience beyond itself), not everything intentional is volitional.[40]

---

[39] The Yogācāra advocate might argue that, according to the doctrine of interdependent origination, insofar as every experiential item invokes *vedanā* (the *skandha* of "feeling" or "feeling tone": positive, negative, or neutral) and every *vedanā* invokes *saṃskāras* (volition), it follows that even simply perceiving a red blotch involves volition. I am not sure, however, that *Arhats* are subject to this, in which case even if it is true of worldlings that color perception invokes volition, it is not true *tout court*, and thus there is no entailment from color perception to volition or, put differently, there is no interdependence between color perception and volition. More importantly, "neutral" feeling tone is arguably just the *absence* of feeling tone.

[40] However, there may even be exceptions to the claim that all volitional states are intentional *because they aim at something beyond themselves*. It seems possible to experience non-intentional volitional states, such as cravings or fears that are unspecific, that is, that lack intentional objects or things at which they aim. Indeed, generalized anxiety disorder ("GAD") seems to be one example that has made it to the Diagnostic Statistical Manual, but there are others. Arguably, even if GAD has no object toward which it aims, it may remain intentional in the weaker sense that it is *about something*, even if what it is about is opaque.

Let me suggest a simple way to preserve—and thus circumvent the conflict between—the differences between the Yogācāra conceptions of causation as karmic (volitional), on the one hand, and the non-Yogācāra Buddhist and Western analytic conceptions of causation as nonkarmic (nonvolitional), on the other. Let us call "volitional^{Y}" (where the superscripted "Y" designates "Yogācāra") the view that all conditioned phenomena and hence all causal phenomena are conception-dependent, thus intentional/volitional, and thus karmic. Importantly, whether or not everything is volitional^{Y}, there remains a valid distinction between those phenomena that are intentional/volitional in the non-Yogācāra Buddhist and Western analytic senses and those phenomena that are not intentional/volitional in the non-Yogācāra Buddhist and Western analytic senses, which latter phenomena we can call "volitional$^{\sim Y}$" (where the superscripted negated "~Y" refers to "non-Yogācāra," and includes other Buddhist and Western conceptions).

This distinction holds regardless of whether the Yogācāra view is correct. That is, if the Yogācāra view is correct, then volitional$^{\sim Y}$ would just be a subset of volitional^{Y}, because not everything that is volitional^{Y} is volitional$^{\sim Y}$, though everything that is volitional$^{\sim Y}$ would be volitional^{Y}. And if the Yogācāra view is incorrect, volitional$^{\sim Y}$ obviously remains valid independently of the category of volitional^{Y} because the Yogācāra view is the only reason on offer to think everything seemingly nonmental is mental.

Thus, it is a separate, open question whether or not everything conditioned or causal is volitional^{Y}. Whether or not everything conditioned or causal is volitional^{Y}, however, it is clear that not everything conditioned or causal is volitional$^{\sim Y}$. Likewise, if we use "karma^{Y}" for the Yogācāra conception of karma, and "karma$^{\sim Y}$" for the non-Yogācāra Buddhist and/or Western analytic conceptions of karma (volition or volitional action), we can say that whether or not everything causal is

karmic^{Y}, it is clear that not everything causal is karmic$^{\sim Y}$. (I will only use these superscripted suffixes when the text requires them for clarity of intended meaning.)

It may be relevant to note that karma may be wiggly in ways in which causation is not. Western philosophers have argued that volitional$^{\sim Y}$ determinism does not function one-to-one because non-volitional$^{\sim Y}$ causes can also affect volitional$^{\sim Y}$ circumstances, just as non-economic causes can also affect economic circumstances.[41] Even within Buddhist thought, volitional$^{\sim Y}$ actions do not invariably lead to precisely the same kind of result, as the present character of a person doing an action also contributes to an effect (AN I 249-53) and future circumstances may affect how and when karmic$^{\sim Y}$ results arise or their degree of intensity. For example, someone who has become an *ārya* can no longer be reborn at less than a human level. But none of this suggests that determinism *simpliciter* runs afoul of one-to-one causal correspondences. Rather, it just enlarges the set of causal conditions to include volition$^{\sim Y}$-extrinsic factors, any set of which together would produce identical—linear, one-to-one—results if collectively repeated.[42]

Shifting gears, Gier and Kjellberg adduce insightful connections between Aristotle's ideas about *the mean* with the Buddha's similar ideas about *the middle path*, apply them to the choices of the mindful versus those of the unmindful, and attribute lack of mindfulness both to ascetic and worldling extremes, extremes that were predominant in the Buddha's milieu. They reason that "the mindful ones are free, while the ones tending to either extreme are not" (284). This insightful observation is in accord with my distinction above between *relatively*-soft and *virtually*-hard determinism, respectively. Further adducing insightful connections

[41] Davidson (*Actions*).

[42] See note 35 regarding multiply complex causal conditions remaining deterministic.

with Western philosophy, Gier and Kjellberg incorporate Keith Lehrer's ideas about meta-preferences,[43] adding:

> We can now define Buddhist free agents as . . . free from ego attachment and craving either for ascetic deficiency or indulgent excess, representing karmic bondage rather than karmic freedom. Free and mindful agents . . . can separate desires from cravings. . . . Buddhists believe in 'free action' but have no conception of 'free will', as a self-determining power that moral agents somehow possess. (284-285)[44]

In applying Lehrer's analysis to mindful volitional cultivation, however, they seem implicitly to accept agents' "self-determining" *compatibilist* power to form and approve volitions. They will acknowledge, shortly, that this conception is plausible within Pāli Buddhism. Their initial resistance to this interpretation owes to the *inflated* construal of autonomy they attribute to the European conception: "It is a supreme irony that what European philosophers assumed is necessary for true human freedom is actually the cause of its greatest bondage" (285). They are referring to the Western idea that the source of free will is in the conception of the self-ruling agent, a conception that has dominated Western philosophy until fairly recently. They admit, however, that if "autonomy" means *not controlled by craving*, then Pāli Buddhists would accept the "concept of agent autonomy, where agents learn to control their cravings and act on their preferences" (285). This volition-

[43] Lehrer ("Preference").

[44] It should be noted that their remarks here about Buddhist free agents would only apply to those Buddhists who are fully enlightened: *Arhats* and Buddhas.

regulating model *would* suffice for a form of Buddhist compatibilism, as I have argued,[45] but they do not develop it.

Instead, they revisit wiggly-determinism (285-86). Referring to David Kalupahana, they say: "For him Buddhist conditionality represents a middle way between strict determinism and the destruction of freedom . . . and an equally unacceptable indeterminism" (285). They characterize Buddhist causation as summed up in the Buddha's conditionality formula, "When that is present, this comes to be . . ." (286), which they equate with *non-linear* interdependence, and they insist that the Buddhist focuses on how to eliminate *duḥkha* (suffering) by eliminating its conditions—something rather *linear* in the twelve *nidānas*. Quoting Edward Conze, they state: "If the total number of conditions is unlimited, and most of them are unknown, it is impossible to say which condition of necessity brings about which event" (287).[46]

In support of this line of thought, one might ask: How can one condition on its own bring about an effect? It seems always to be the case that several conditions are needed, with some of these being necessary conditions. Often what we mean by "the" cause of something is the last condition to fall into place, or perhaps the most significant one. As Buddhaghosa, speaking on the twelve *nidānas* (in the *Visuddhimagga*), says, "there is no single or multiple fruit . . . from a single cause; nor a single fruit from multiple causes. . . . But one representative cause and fruit are given in this way, 'with spiritual ignorance as condition are the constructing activities'" (*Vism.* XVII 106).

Notwithstanding Conze's and Buddhaghosa's intuitive claims, complexity and opacity do not entail non-linearity, as I argued above.

---

[45] Repetti ("Meditation").

[46] Conze (146 fn.).

Indeed, the most threatening version of hard determinism focuses on maximally-complex *world-states*, such that the entire world-state W_1 at time T_1 determines W_2 at T_2 in accordance with all the laws of the universe (say, L_1, L_2, L_3, . . . L_n), collectively, "L": W_1 nomologically-necessitates (or L-necessitates) W_2, W_2 L-necessitates W_3, and so on, and we are powerless to alter this globally-determined sequence.[47] But the nomological elements in global-states do not *invalidate* the linear sequencings that form those global-states: rather, because global world states and their local constituents are *interdependent*, the whole "interdependent *versus linear*" dichotomy is a false dichotomy, easily dismissed by describing interdependence as *multi-linear*.

There is another problem with the issue of the causation of world states. Most Buddhists accept mereological (part/whole) reductionism:[48] any apparent whole that decomposes to a partite aggregate (a collection of parts) is a mereological fiction. This is classically illustrated in the *Milindapañhā* with Nāgasena's chariot example: a chariot is nothing above the configuration of parts designated as such relative to our interests. One may object, therefore, that "global" or "world" are the *greatest* aggregate-based mereological fictions; and because determinism and interdependent co-origination (conditionality) are both *global* in scope, from the mereological reductionist perspective to which Buddhism is committed, both determinism and conditionality are, on analysis, mereological fictions!

The Pāli material, one might reply, focuses more on a small number of conditions—say, twelve (the twelve *nidānas*)—for any state, in which case the Pāli theorist might reject the global interpretation of dependent origination. However, both mereological reductionism and

[47] See van Inwagen (*Essay*).

[48] Siderits (*Persons*).

the no-self doctrine extend to all conditioned phenomena, which are thus "empty"—lacking inherent natures, independently existing essences, or ontologically objective lines of demarcation that could individuate entities (parts/wholes) in ultimate reality. Certain *Abhidharma* atomists who assert that at the ultimate reality level there are impartite psychophysical atoms are an exception.[49] However, being able to individuate psychophysical atoms is not the same as being able to block laws about their aggregation-level behavior in the form of world-states (or, on a much smaller scale, in the form of mental states), at least not at the level of ultimate reality, where everything is interdependent. If everything is interdependent, then nothing is independent of anything else, and thus every putative atom is really just an aspect of the globally interconnected totality.[50]

Nonetheless, suppose we reject "global" but retain the nomological connections between all interdependent conditions (C_1, C_2, . . . C_n),[51] collectively, "C," that jointly constitute the world-aggregate W_1 at time T_1 and jointly cause event E_1 under L at T_2. But all events (E_1, E_2, . . . E_n), collectively, "E," at T_2 that constitute aggregate-W_2 are caused by C. So, C produced E (aggregate-W_1 caused aggregate-W_2), which means C_1, C_2, and so on, simultaneously produced E_1, E_2, and so on. Even so, this analysis does not obliterate the linear determinism that obtains between aggregate-W_1 and aggregate-W_2 (such that it is a law that *W_1's cause W_2's*) or

[49] For a relevant discussion of *Abhidharma* atomism, see Siderits (*Persons*).

[50] In fairness, this line of objection ought to be directed more at Buddhism than at Gier and Kjellberg's attempt to articulate it.

[51] These may be psychophysically atomic conditions, if *Abhidharma* is correct, or quantum potentialities, if physics is correct, or both, if both are accurate descriptions of the same phenomena under different modes of presentation, as an East/West convergence metaphysician like B. Alan Wallace might put it (*Dimensions*). For purposes of the line of reasoning in the text, however, it does not matter which interpretation is correct.

that might still obtain between each element of C and each element of E (for example, between C_1 and E_1, C_2 and E_2, and so on). The same holds for the indefinitely-many cause/effect pairs that constitute any two successive aggregate-world-states,[52] even if every psychophysical atom or quantum-energy-unit causes effects in, and is causally affected by, every other psychophysical atom or quantum-energy-unit in the same instant. Thus, interdependence makes possible indefinitely-many more linear causal pairs for determinism—yielding a *mega-linear*-determinism![53]

Moving away from the issue of causation, Gier and Kjellberg turn to the Pāli conception of the self as a *process*. As we shall see when we turn to our review of the accounts of Asaf Federman and of Peter Harvey, most scholars who see the Pāli conception of the self as a mere process rather than as an entity consider the Pāli conception to be deflationary. However, Gier and Kjellberg's characterization seems somewhat inflated, for they say that the Pāli self is depicted as "a robust personal agent fully capable of maintaining its personal integrity and taking full responsibility for its actions" (289). Their implicit support for this somewhat inflated description seems to contrast with a more minimalistic conception of the self they will describe shortly in connection with Mahāyāna Buddhism.

---

[52] However, recalling the Buddhaghosa quote in the text above, dependent origination is not about one thing, on its own, being the cause of another. For example, although it is said that craving is dependent on feeling, this only applies to unenlightened beings that still have ignorance. Thus, the same putative cause does not have the same effect (in beings at different stages along the path). All things considered, if dependent origination is not deterministic, then there simply is no problem in Buddhism analogous to that posed by determinism for free will in Western philosophy. If dependent origination is either not deterministic or not coextensive with determinism, and determinism is true, then Buddhism simply has to come to grips with it.

[53] This criticism, too, is not directed specifically at Gier and Kjellberg, but rather to any Buddhists who embrace a wiggly determinism.

In further dissecting the idea of a free agent, they deny a Buddhist notion of "the will" but emphasize *cetanā* (volition). They claim the Pāli conception of the person involves a middle-way posit between nihilist (*anātman*: the negation of the existence of a self) and eternalist (*ātman*: soul, conceived as eternal) extremes, and they conclude that Pāli Buddhism "falls under the general rubric of compatibilism and 'soft' determinism" (289). I agree with many of their claims, but their supports for these claims are mostly scattered and inexplicit.

In the remainder of their analysis, Gier and Kjellberg focus on Mahāyāna sources, mainly Nāgārjuna. After outlining three standard arguments (impermanence, reductionism, and insubstantiality) for the insubstantial view of the self, they paraphrase Nāgārjuna, stating "if we cannot say that the self really exists, by the same token we cannot say that the self really does not exist" (291). They do not explain here why it would be a mistake to assert or deny the self, though in their subsequent discussion they seem to be laying out an understanding of Nāgārjuna's thought on the basis of which the reader may surmise the explanation. Based only on their presentation, however, that understanding is not obvious, although this difficulty likely owes more to Nāgārjuna's ideas than to Gier and Kjellberg's attempts to explicate them. Instead, Gier and Kjellberg problematize the very Pāli/Mahāyāna distinction that is the target of their work:

> The Pali versus Mahayana distinction is now not very helpful, so henceforth we will distinguish between 'constructive postmodernism' (CPM) and a skeptical 'deconstructive postmodernism' (DPM). . . . CPM is generally realist and supportive of the canons of logic and evidence, while DPM rejects realism and any logocentric methodology. Nagarjuna is a consummate logician and

> never rejects logic as a standard, so this obviously causes problems for any DPM reading of him. (291-292)

Although this somewhat anachronistic juxtaposition is insightful, and although these differences promise to be explanatory regarding the Buddhist free will issue, Gier and Kjellberg draw from them no *explicit* conclusion about Buddhism and free will. This supports my impression, suggested earlier, that they are simply surveying important elements of terrain that need to be charted in order to fully explicate the Buddhist understanding of free will. This seems helpful, as far as it goes.[54]

On a more interesting note, they suggest "that the phrase 'phenomenal self' be used for the Pali tradition and CPM interpretation while reserving 'conventional' as a placeholder term for the deluded self who thinks it lives in a real world of interdependent things and events" (292).[55] This implies that there is a three-truth-levels division consisting of: (1) ultimate; (2) conventional, the relatively deluded perspective of the worldling; and (3) an intermediate reality level, the relatively non-deluded perspective of the *ārya*.[56] After raising this trifurcation, they claim that because Pāli Buddhists have a realist concept of truth, they can embrace compatibilism, but the explanation for this claim, like many

---

[54] Recall that their article was written as the only non-Western piece for an edited collection of otherwise all Western views on free will.

[55] Presumably, they meant not "interdependent" but "independent." It should be noted that their use of "conventional" here is somewhat non-representative.

[56] This trifurcation resembles the Pudgalavādin (person-affirming) view (Priestly, "Pudgalavāda"), but is arguably equivalent to a two-fold ultimate/conventional parsing where the conventional divides in two; there may be valid or invalid conventional views. It is uncertain how many Pudgalavādins there have been overall, although they constituted roughly a quarter of the Indian *Saṅgha* (the Buddhist monastic community) around the time of the conventions that led to the inscribing of the Pāli Canon, so the casual dismissal of their interpretation among many contemporary exegetes is not necessarily warranted.

others, is left to the reader to surmise from their subsequent remarks—in this case remarks about Mahāyāna anti-realism that implicitly contrast with Pāli realism.

For example, they suggest that if the correct view of self is the Mahāyānist's thoroughly-conventional self (recall, they reserve the term "conventional" for the deluded self),[57] "then to ask whether it is free or determined is like asking 'what is the sound of one hand clapping?'" (293). However, not all illusions or falsehoods are equal, nor are they equally mysterious or paradoxical. According to such diverging theorists as Peter Harvey and Mark Siderits,[58] the "deluded self" exists as much as anything else, though its deluded projection posits a non-existing thing.

Further sketching important elements of terrain, Gier and Kjellberg discuss Nāgārjuna's assertion of the interdependence of agent and action. Nāgārjuna's idea is that there cannot be an agent without an action or a potter without pot-making: the items in these pairs are interdependent and cannot exist independently of their other pair members.

---

[57] Does this mean that to believe in self is a thorough delusion? They do not say, but this is not obviously a pan-Mahāyānist view of the self.

[58] Harvey (communication, April 2013); Siderits (*Persons*). However, as Dan Cozort has emphasized (communication, July 2013), this interpretation is not shared among all Mādhyamikas (followers of Mādhyamaka, the Mahāyāna "middle way" philosophy, first clearly articulated by Nāgārjuna, between the extremes of the eternalism of *ātman* and the nihilism of *anātman*). Thus, although Bhavaviveka (a Svātantrika-Mādhyamika) would say that an inherently existent self exists conventionally, Candrakīrti (a Prāsaṅgika-Mādhyamika) and his followers would say that it does not, because such a self has no basis of designation. The example of the rope mistakenly taken for a snake illustrates for Bhavaviveka that the rope-snake exists conventionally, but for Candrakīrti, there is no rope-snake, just a rope mistaken for a snake. Nothing about a rope, on his view, could be the basis of designation for a snake. Given these divergent interpretations that bear importantly on the understanding of the self, there is no "Buddhist view" of the self.

However, one may object that there is an unequal direction to the interdependence: *pot-making* does not *make* the person a potter, even if *in the process of pot-making* the person *expresses* or *manifests* that she is—and, perhaps if for the first time, she thereby *becomes*—a pot-maker. The person is a potter only when making a pot.[59]

In concluding this agent/action discussion, Gier and Kjellberg state: "Normally we assume there has to be a self or agent in order for there to be freedom, but this is just the presumption the skeptical Nagarjuna questions" (297). Is this meant to imply that they take Nāgārjuna to solve the free will issue, or to dissolve it? This is an important question, but they do not even raise it explicitly, much less suggest an answer to it. Instead, they simply problematize the issue, raising further complications regarding the self and the no-self views, existence and non-existence, the four-cornered negation, and so on—all suggestive of but lacking explicit explanation, and all supporting my view that they are simply identifying terrain rich in explanatory potential.

For example, Gier and Kjellberg state: "If we cannot call the karmic web free since it lacks a self, by the same token we cannot call it determined, since nothing outside of it is causing it" (297). But the opposite conclusion makes more sense: without *agency* in the karmic web, the *only* thing that could explain web-activity is that it is, or its components are, determined. This raises the question whether their reasoning about self-lacking karmic webs applies to enlightened persons, who are in some sense self-lacking. It seems that agentless enlightened persons might still have karmic "webs," even though they are not *entangled* in

[59] This objection resembles a similar objection about the direction of explanation that Socrates pressed on Euthyphro about the latter's idea that piety is *whatever is loved by the gods*. Socrates seemed to think that an action's being loved by the gods does not explain *why* the gods love the action, but rather that the gods love the action *because it already is pious*, even if it is correct that *whatever is pious is also thereby loved by the gods*.

them and even if "their" past karma continues to unfold; only ego-volitional unenlightened beings are web-entangled.

One *could* say that the web of conditions that constitutes an enlightened being lacks a separate self or a deluded sense of one. Because that being/web is in form and all form is conditioned, that being/web is causally determined (karmically or not). But nothing that determines it does so in a *freedom-undermining* manner, at least not on certain compatibilist models according to which the absence of freedom-undermining influences plays a constitutive role in ascriptions of free will. According to Harry Frankfurt, for example, weakness or failure of will is impossible without a separate meta-volitional system in which there is meta-volition/action discord—for example, when the agent identifies with a metavolition to the effect that a particular volition should succeed in governing or leading to action, but the meta-volitionally-approved volition fails to do so.[60] My impression of *nirvāṇa* is that because it entails the extinction of the *Arhat*'s identification with volitions, it arguably eliminates the possibility of meta-volitional/volitional discord. It is not clear that enlightened persons lack volitions altogether, for it seems plausible that they experience physiological states like thirst, aversion to physical pain, and the like, and in their boundless altruism and compassion they seem to intend the welfare of all sentient beings. Presumably, however, they are not attached to any of these intentions. Indeed, as

[60] Frankfurt ("Freedom"). Some interpretations of enlightened beings move in the opposite direction. For example, according to Peter Harvey, enlightened beings are said to have a "great" or "developed" self, that is, a strong, calm, self-contained personality. See Harvey (*Selfless*), chapter 3 and (*Buddhism*), 62. According to Harvey B. Aronson (*Practice*), Buddhism denies an *ontological* self, not a *psychological* self, and seeks to destroy the delusion of an ontological self while strengthening the psychological self. (I owe this observation to Dan Cozort in a communication, July 2013.) This interpretation resonates well with Harvey's account ("Freedom"), Federman's account ("Buddha"), and my own account ("Meditation").

Christopher Gowans has argued, they are not even attached to the Dharma (*Philosophy* 59-60).

Although it is not clear whether Buddhism would accept this possibility, consider for the sake of argument that there are certain unenlightened beings such as insects and other primitive life forms that constitutionally seem to lack any sort of self-sense because they lack sufficiently complex mental states.[61] These creatures are completely stimulus/response determined by *exogenous* conditions impinging upon their *endogenous* conditions in accordance with natural laws. If there is no self-sense, they *must* be exogenously determined. Humans, except

[61] One reason Buddhists might resist this possibility is that all unenlightened sentient beings are thought to experience hedonic states and to be conditioned by ignorance, attraction, and aversion—primary conditions of the self-sense. It is not clear that the Buddhist conception of the self-sense requires an explicit, conscious, or Cartesian thought of the form "I am," as, perhaps, an anthropocentric animal psychologist might require. Another possible source of resistance might be the widespread Buddhist belief that non-humans lack the sort of full-fledged intentions that define typical human mental states and that are necessary to accumulate karma. My intuitions run counter to this idea because it seems that although instinctive or not-fully-conscious *behaviors* are possible without intentions, no *actions* are possible without some intention, and many nonhuman animals engage in what appears to be teleological or goal-oriented and thus intentional behavior—actions. I cannot imagine how animal reincarnation is possible without karma (and all karma—action—is intentional). But it strikes me as though the issue whether nonhumans have volitions is independent of the issue of their having any kind of self-sense whatsoever, as it seems intuitive that there may be some sort of rudimentary sense of self in any sentient beings that are pain-aversive. For evidence and argument in support of the view that all vertebrates and many if not most invertebrates have intentions and rudimentary self-sense, see DeGrazie (*Animal*), chapters 3 and 4. The point in this line of thought is that although the description of enlightened beings as *sentient beings lacking a self-sense* presumably picks out something true of enlightened beings, insofar as *sentient beings lacking a self-sense* conceivably includes unenlightened beings as well, this description is too inclusive and thus inadequate as a definition.

perhaps those who are enlightened, have elaborate self-sense systems.[62] Gier and Kjellberg try to explain this asymmetry:

> To the extent that people identify a self, that self is determined by causes outside of it. The more cultivated they become on the Buddhist model, the less they think this way. The less who think this way? A question that the European philosopher might ask. Nagarjuna's answer is, no one, really. The nonpersonal web of causes and conditions sheds the delusion, or, rather, ceases to give rise to it. (297)[63]

However, the mystery of how impersonal webs both cause and dissolve delusions needs an explanation. Gesturing toward an explanation, Gier and Kjellberg continue:

> Thus, you get the seemingly paradoxical lines from, for instance, the Diamond Sutra, that 'even though infinite beings have been saved, none have been saved.' Thus while we would assume that there has to be a self in order

[62] The reason for my cautious tone is that if Aronson is right in his claim that enlightened beings do not lack a psychological sense of self but only lack the illusion that this integrated set of psychological functions constitutes, reflects, or manifests an ontological self (*Buddhist Practice*), however, then the descriptor, *sentient beings lacking a self-sense*, would not even apply to enlightened beings.

[63] However, Gier and Kjellberg are not maintaining their own distinctions, for if, as they say earlier, the "self" they mention in their first sentence in this quote is a delusion, then there is no "it" to be determined. It should be emphasized that although it is a philosophically legitimate question whether animals have intentions, karma, and/or a self-sense, as suggested earlier Buddhists mostly have assumed non-humans do not have the ability to form intention. The Jatakas have many stories of animals performing virtuous acts, but these are standardly taken not literally, but metaphorically.

> for there to be freedom, Nagarjuna would say that there is freedom only to the extent that there is not a self. (297)

As astute as is their identification and analysis of this problematic dimension of the issue, it rests on a simplistic equation of the lack of a self-sense with liberation, on the one hand, and it ignores the crucial distinction between psychological and ontological senses of the self, on the other hand.

Gier and Kjellberg suggest that the presence of a self-sense is what constitutes bondage, and its absence freedom, but if *nobody* ultimately exists, who experiences the self-sense? Who is bound? It may be objected that the Diamond Sutra remark *states* or *embraces* the wasp/Buddha puzzle, but does not *explain* or *resolve* it. Applying Aronson's distinction between psychological and ontological conceptions of self (*Practice*), I have hinted at a way to explain the Diamond Sutra remark and the wasp/Buddha puzzle by reference to the presence of a delusional sense of an ontological self in the unenlightened and the absence of a delusional sense of an ontological self in the enlightened, together with the presence of a functioning psychological self—the highly-integrated functionality of the five *skandhas* ("aggregates": form, sensation, perception, volition, and consciousness)—in both the unenlightened and the enlightened, but this is only a sketch. On this interpretation, "the "I am' conceit" is the delusional sense of an ontological self. Nonetheless, this is a deeply puzzling issue, so it is no reflection on Gier and Kjellberg that they do not explicitly resolve the puzzle, but instead simply point toward its solution.

The issue may be thought to boil down to an empirical question: whether creatures that only experience pleasure and pain and attraction and aversion thereof, respectively, have "the 'I am' conceit." Arguably, "the 'I am' conceit" requires something on the level of a Cartesian thought, but it is doubtful all merely sentient beings have the requisite

psychological complexity. It is possible that some sentient beings even lack hedonic states at all, despite their ability to detect changes in their organism/environment field. Gier and Kjellberg claim that the principal culprit in suffering is *belief* in a *real* self (*ātman*), which belief "gives rise to desires that are ultimately unsatisfiable precisely because no such self exists" (297). Conceivably, this complex way of conceptualizing the issue might diverge from something much more primitive: there may be a conceptually simple or pre-conceptual form of a delusional sense of self. But *belief in a kind of self-nature* is a sophisticated psychological achievement, however ill-conceived it may be—an achievement presumably beyond the cognitive scope of many species of sentient beings.

This critical line of thought, too, aims more at Buddhism per se than at Gier and Kjellberg's attempts to explain it. Nevertheless, a non-Buddhist might ask, how does a non-entity (like phlogiston) generate beliefs—that it, a *nobody*, somehow *believes*—that generate *desires* that belong *only to the non-entity* (as opposed to another non-entity such as the Buddha), without an intentional agent, as per Nāgārjuna's insistence about *no action without an agent*? What may be said on behalf of the Buddhist view here (which Gier and Kjellberg do not articulate) is that there is no real self, but the *skandhas*—within which the belief arises (and in connection with which there are volitional and related implications and consequences)—are not non-entities. The Buddha's exposition about dependent origination—in particular, about the way the twelve *nidānas* operate—is meant to suffice here: these ultimately impersonal factors operate in such a way that the illusory self-sense arises, grasps, and so forth.

Perhaps implying that the self-issue has been adequately problematized, Gier and Kjellberg shift focus and claim that because the aim of Buddhism is to *eliminate* the self, the Mahāyānist rejects *personal* enlightenment because that *reifies* the self. They elevate the ideal of the

bodhisattva, one who seeks enlightenment for the sake of others (298). But the links in this line of their reasoning are somewhat cryptic, so let us pause to reflect on the idea of eliminating the self, on one hand, the idea of seeking enlightenment for the sake of others, on the other hand, and on the idea that these are supposed to be related as well.

First, let us discuss the idea of eliminating the self. The way Gier and Kjellberg discuss the elimination *of the self* is somewhat misleading, as we have already seen. In Buddhism a key aim is to directly know that everything is non-self (not a permanent, substantial self), and this insight is taken to lead to the elimination *of the deluded view* of being such a self, and to the elimination of "the 'I am' conceit" and its effects: self-importance, self-centeredness, ego-volition, and so on. This differs from the way Gier and Kjellberg speak—as if there is a self and it is to be eliminated.

Let us now discuss their bodhisattva claim, and how this relates to their idea of eliminating the self. If bodhisattvas—or *āryas* or worldlings, for that matter—are ultimately non-autonomous non-entities, a skeptic may ask how can "they" *bring it about* that they attain or *control* anything for anyone's sake? Whether one is a worldling, *ārya*, or bodhisattva doesn't seem to make a difference here. Arguably, paradoxical wisdom *implicit* in the Diamond Sutra *could* explain these apparent contradictions, but as presented by Gier and Kjellberg they remain mere hints of transcendent wisdom.

Let us try to further spell out what they seem to imply. As Aronson has made clear (*Buddhist Practice*), Buddhism differentiates between *ontological* and *psychological* models of the self, and it denies the existence of a certain type of ontological self, defined in various ways (substantial, inherently existent, unchanging, essence-bearing, independent, and so on), but it does not deny the psychophysical processes that generally constitute us—it does not deny self *tout court*.

Those ultimately impersonal psychophysical processes—the *skandhas*—are erroneously taken to be indicative of an ontological self, and it is that false consciousness that vanishes upon enlightenment. The bodhisattva ideal is thought by some to be soteriologically superior insofar as its altruistic orientation renders it less likely to motivate practitioners by a false ontological sense of self that might be engendered by seeking the end of one's own suffering. Most Buddhists who favor both approaches would agree that enlightenment is attained when the false ontological conception of self is eliminated—not when the psychological self is eliminated.

Because the crucial difference between enlightened and unenlightened beings is, on my analysis, the absence or presence of belief in a false ontological sense of self, and the object of that belief—the ontological self—does not exist, although beings attain enlightenment when the false ontological sense of self is eliminated, in ultimate reality there are no ontological selves who attain enlightenment. Because actions occur without any ontological selves to author them, there is no contradiction in the idea that practitioners—who are, in ultimate reality, nothing over and above the collection of the *skandhas*—can "do" things that lead to their enlightenment.

Gier and Kjellberg do a fine job of surveying the many problematic and puzzling terrains of Buddhist doctrine and its divergent interpretations that need charting by any cartographer of the Buddhist understanding of free will. They seem to conclude that Pāli Buddhism is compatibilist but naïve, and that Mahāyāna Buddhism transcends the free will issue and linear determinism, and is more plausible; but many of their more detailed claims are sketchy and many of their more pointed arguments are inconclusively suggestive. But they do importantly introduce and emphasize major differences in view among Pāli and Mahāyāna Buddhist philosophers.

Let us turn now to review how scholars on both sides of the Pāli/Mahāyāna divide interpret the issues Gier and Kjellberg discuss. Let us begin with the earlier, Pāli views, as expressed by Asaf Federman and Peter Harvey, beginning with Federman's account.

Federman: The Buddha *Taught* (Theravādin) Compatibilism?

Asaf Federman's article is entitled, "What Kind of Free Will Did the Buddha Teach?"[64] Of course, because the Buddha never explicitly discussed "free will" or "determinism," technically he never *taught* any kind of "free will." Citing many canonical passages, Federman argues that the Buddha had a "deflationary" conception of volition such as is found in the thinking of Daniel Dennett, and he concludes, basically, that Dennett-style compatibilism is "what the Buddha taught."

Federman first differentiates "inflated" and "deflated" conceptions of agency. His "inflated" model involves a contracausal, acausal, or transcendent ability that exempts agency from the physical domain and that he considers *inconsistent* with determinism. The problem with this view, however, is that it makes light the Buddhist acceptance of things nonphysical: the inconsistency is with the *physicalism* implicit in the standard interpretation of determinism, not with determinism per se, for insofar as Buddhism avers that dependent origination applies to all conditioned phenomena and that some conditioned phenomena are nonphysical, it leaves open the possibility that causality applies to the nonphysical. Federman's "deflated" model involves a natural ability to

[64] Published in 2010, but a draft was presented at a 2007 conference. Wallace refers to an online version of the paper as a later version (relative to the conference). I review Federman's 2010 publication prior to Harvey's 2007 or Wallace's 2008 publications because both reference Federman's 2007 draft.

reflect on desires, make intelligent choices, act accordingly, and evade undesirable outcomes, and that he considers *consistent* with determinism. The problem with this view, conversely, is that it makes light of the Buddhist view of the ultimate insubstantiality of the self or agent and thus, arguably, is not deflated *enough*. Because Federman is advocating for a Theravāda-based view here, perhaps this is what Gier and Kjellberg had in mind when they stated that the Pāli self is depicted as "a robust personal agent fully capable of maintaining its personal integrity and taking full responsibility for its actions" ("Buddhism" 289).

Federman offers two corresponding definitions of free will. The inflated version "defines free will as a power that belongs in the soul, transcends the physical, and has ultimate control over the body" (3).[65] Although this sounds very much like *libertarianism*, Federman does not mention this standard term here. The deflated version "defines free will as the agent's ability to control action in conformity with will, when there are no constraints that limit performance" (3). Although this sounds very much like *soft determinism*, Federman does not use this standard term either.

Let us call these two models of free will "inflated-FW" and "deflated-FW." A problem with Federman's eschewing inflated-FW is that because the Buddha seems committed to the view that non-physical mental factors causally control actions, physical movements and the like, Buddhism would not obviously be required to reject inflated-FW *on that ground alone*. From a Buddhist perspective, that something is non-physical does not entail that it exists outside the causal nexus. Buddhism *would* reject the idea that non-physical mental factors that causally

[65] By "transcends the physical," Federman means it is immaterial in the Cartesian sense. However, not all libertarians are dualists or immaterialists. See, for example, Kane ("Pathways").

control physical movements and the like inhere in an ontologically substantive, independently existing, essence-possessing soul, but it would be open to accept a deflated or minimalistic conception of the non-physical psychological processes that constitute the psychological self and to attribute *to them* a causal control over actions and certain bodily behaviors associated with them, which analysis might equate with a minimalistic form of libertarianism or partly-inflated-FW. However, it is difficult *for non-Buddhists* to take seriously the idea that the non-physical remains causal. For example, monists reject dualism because it is exceedingly difficult to imagine how anything non-physical like a mind or soul could causally interact with anything physical like a brain or body.

A potential problem with endorsing deflated-FW is that very young children and even wild animals appear to have it, for they can choose where to run. Harry Frankfurt characterizes this (deflated-FW) as *freedom of action* rather than *freedom of the will*, the latter of which involves volitional regulation ("Freedom"). On his analysis, for example, an animal can run in the direction it desires to run, *controlling its action in conformity with its will*, but is not an "agent" (in the moral-responsibility-entailing sense at issue in the Western discussion of free will), nor does it exhibit freedom of the will unless it has the sort of will it wants to have—that is, its effective desires (those that succeed in leading to actions) accord with its meta-volitions (its higher-order volitional preferences). Thus, if deflated-FW is mere freedom of action, then it is over-inclusive, too weak.

Another problem with deflated-FW involves the no-self doctrine. If there is no such thing as an *agent*, then Federman's description, "the agent's ability to control action in conformity with will," is too strong. This raises the problematic question whether "impersonal agency" may be explicated coherently under Federman's Theravādin interpretation of

Buddhism. A thermostat is no agent, but it exhibits *proximal* control over a heating system: set to 72°, the thermostat controls the states of the heating system to maintain 72°, but lacks free will because it is unable to alter its own settings—that is, it is unable to bring about what Frankfurt might describe as *having the temperature setting it wants to have* (analogous to having the will it wants to have), as opposed to *controlling the heating system in accordance with its temperature setting* (analogous to Federman's criterion of controlling action in conformity with will). Its setting is not "up to" it, so to speak, but freedom of the will entails that our choosing from among our desires which ones we approve of and act on is "up to" us in some important—even if inchoate—sense.

Federman says that the "more accurately one represents reality and imagines possibilities, the more freedom one has" (9), but a Wi-Fi cyborg-thermostat that can detect subtle variances in temperature throughout a sky-riser and analyze online a vast array of possible sub-routines regarding the most efficient blend of solar, oil, and gas heating alternatives relative to current and projected market prices would therefore have more "free will" than an ordinary thermostat. In our case *more* of the command chain is located within our brains (we alter our own settings, in some sense), but in both cases scientists and Buddhists agree that all causal command chains are ultimately exogenous and impersonal, and nothing in our behavioral system exhibits *ultimate endogenous* control, what might be described as our volitional lives being *ultimately up to us*. Nonetheless, the snail and the thermostat satisfy deflated-FW. Thus, inflated-FW is too strong, and deflated-FW is too weak.

These problematic edges of the discussion aside, Federman's main argument is sound, boiling down to the following four premises, illustrating astute canonical excavation and exegesis:

1. The Buddha rejected the idea that we exist outside the causal nexus (6);

2. The Buddha rejected the idea that the will is impotent (9);

3. The Buddha advocated that by making the right choices, we can progress toward enlightenment (10); and

4. The Buddha asserted that everything is dependently originated (11).

These four premises lead to Federman's implied conclusion:

5. Therefore, the Buddha taught compatibilism (deflated-FW) (*passim*).

Federman's argument for a Buddhist theory of free will is one of the most plausible among those I have reviewed,[66] despite those weaknesses noted above. For instance, assuming the Buddha held 1-4 does not guarantee that he held or "taught" what Federman asserts in 5. To insist that *because one holds beliefs P and Q* one must also *believe that P and Q are compatible* is to commit Barnhart's fallacy.[67] If 5 were expressed more cautiously, for example, "What the Buddha taught is *consistent with* compatibilism," then Federman's argument would be much stronger.

In his brief review of Mark Siderits's proposals, Federman scrutinizes his two truths approach, stating "Although he argues that in Buddhism personal freedom and psychological determinism relate to each other like 'two ships passing in the night,' he admits that the Buddhist rejection of *ātman* practically cancels the possibility of free will" (2).[68]

---

[66] See Repetti ("Early"; "Reductionism"; "Hard Determinism").

[67] Named after Michael Barnhart, who noted (in conversation) that Asian philosophers sometimes exhibit "a blissful maintenance of contradictory beliefs," in which case it is fallacious to assume they *believe* they are compatible. See Repetti ("Reductionism").

[68] The embedded quote is from Siderits ("Beyond"). See Repetti ("Reductionism") for a critical review of Siderits's complex attempt to use the two truths doctrine to parse

Federman reasons that, "although the Buddhist doctrine of not-soul rejects the idea of ultimate self-control, this does not lead to denying that people control their behavior and choose their actions" (2). However, Siderits does not deny choice, but only that persons exist *ultimately*. The deflated-FW view—that anyone who makes choices, does what they want to do, and exhibits proximal control over what they do (for instance, they raise their arm when they want to) has free will—appears to be a form of "naïve autonomism," which on analysis very much resembles Frankfurt's *freedom of action* (as opposed to *freedom of the will*, the meta-volitional ability to have the will one wants to have).[69] Rats, snails, and thermostats exhibit elements of naïve autonomism; therefore, deflated-FW is not sufficiently fine-grained to represent the sort of free will—in the moral-responsibility-entailing sense—that we have in mind in philosophical discussions of free will.

In his brief review of Gier and Kjellberg's proposals, Federman says that their work "leads to an unsatisfying conclusion: that Buddhism is silent about free will because its conceptual toolkit is different from the modern toolkit" (2). However, the toolkit claim is only in Gier and Kjellberg's introduction. Federman ignores the bulk of the groundwork they lay for the very Pāli compatibilism he espouses, significant elements of which I have reviewed above.

---

libertarian free will as conventional and determinism as ultimate, claiming that there is no self at the level at which determinism holds.

[69] Frankfurt ("Freedom"). Although it is true that Frankfurt is much more well known for his article rejecting hard determinism ("Alternate") than he is for his article sketching his own model of the implied default (compatibilist) position thereafter ("Freedom"), it is, to my lights, a major oversight that few scholars have taken seriously his identification of the relative importance of freedom of the will versus the relative unimportance of freedom of action, as even my few remarks about this distinction here seem to amply illustrate.

Turning to his own analysis, Federman says that the idea that a Cartesian "non-physical agent imposes freedom on a closed causal system contradicts essential aspects of scientific thought" (5), but he ignores canonical passages that make reference to contracausality, psychic powers, astral worlds, and related Buddhist ideas about the existence of non-physical items (Kalupahana 21-23, 40-44, 114-115) that likewise seem *prima facie* to run afoul of "a closed causal system" and equally "contradict essential elements of scientific thought."[70] The idea that non-physical mental conditions could have causal influences on physical items might violate the physicalist's principle of the closure of the physical domain, but it does not violate any principle of Buddhist thought. Arguably, Buddhism accepts a principle of closure of the psychophysical domain. Thus, the idea that a Cartesian ego—which Federman rightly likens unto an *ātman*—could influence matter is not antithetical to Buddhism *on that ground*, but only on the ground that Buddhism rejects the notion of that particular type of non-physical entity.

Federman says, "Unlike some libertarian positions that emphasize only the freedom of action, Buddhist liberty primarily refers to the mind being free from what binds it" (10). There *is* an important distinction between Buddhist *mental* freedom and freedom *of action*, but libertarianism does not emphasize freedom of action—rather, on my analysis above, Federman's deflated-FW does. An irony here is that some libertarians believe they possess freedom of the (non-physical) *mind* from the bonds of (physical) determinism in a way that could be consistent with what I suggested above about the Buddhist acceptance of a principle of

[70] Buddhism arguably sees all these as arising according to conditions as well, but whether Buddhism holds that *nonphysical* conditions are subject to *the same sorts* of causal laws or mechanisms is not clear, and whether or not this undermines the idea that the doctrine of dependent origination is to be construed as deterministic seems problematic at least, and an remains an open question at best.

closure for the psychophysical domain. When the Buddha converses with celestial beings (SN I 1, I 23-31, II 6, II 17), for example, this suggests *mind* can operate *outside* the physical domain.

Standard Buddhist arguments against the coherence of the notion of a causally inert *ātman* ground the idea that there is a distinction in Buddhism between causally potent psychic phenomena, such as *saṃskāras*, and the causally inert soul. Of course, a Buddhist can *claim* that dependent origination—however it is to be understood—extends to every dimension of conditioned phenomena, heaven worlds and the like included. But whether it makes sense for non-Buddhists to think that laws determine non-physical things—when *our* model for laws is that of necessary relations between physical states—is another question.

One may object that most of the *nidānas* (causal links) of dependent origination are mental in nature, but this seems only to highlight the need for an account of the nature of mental causation clear enough to know whether or not it is simultaneously physical and deterministic. Intuitively, it seems not to be, if there are celestial beings, or if reincarnation is not necessarily simultaneous with death, such that beings exist in non-physical (post-mortem, pre-reincarnational) realms. Of course, if all that exists are aggregates of psychophysical atoms, as *Abhidharma* Buddhism posits, then perhaps they can function equally on a non-physical level that is nonetheless causal, just as they apparently do on the physical level, given their two-fold natures.

Although it is worthwhile to note these problems with Buddhist metaphysics and their possible solutions, we cannot resolve them here. Instead, let's characterize Buddhist *mental* freedom in *negative* terms by the *absence* of mental bondage or constraint, the maximum satisfaction of which is in the enlightened being's *liberation* in *nirvāṇa*; let's call this kind of *nirvāṇa*-based freedom "*nirvāṇa*-F," in line with our abbreviations

for deflated free will ("deflated-FW") and inflated free will ("inflated-FW"). Federman discusses the case of the otherwise spiritually evolved bodhisattva—presumably an *ārya*—who satisfies deflated-FW in deliberately performing an adharmic sacrifice,[71] but because he is in the grip of romantic passion, lacks *nirvāṇa*-F. However, not only the *ārya*, but also the worldling often satisfies deflated-FW without satisfying *nirvāṇa*-F, and surely the Buddha *actually* "taught" *nirvāṇa*-F and its absence.

Federman also advocates wiggly-determinism, supposing "dependent-arising is not a linear causal sequence" (12), and juxtaposes karma and determinism, stating "there are no strict deterministic relations between the act and the result" (13). By in engaging in this line of reasoning, however, Federman seems forgetful of his own resistance to things that violate physical law. Federman correctly notes that determinism and fatalism differ importantly and that their conflation infects the discussion; he describes fatalism as "an ethical stance that states that choice is meaningless" (13), but he is insufficiently clear about the difference. I have differentiated between (hard, soft, and wiggly) determinism, karma, and fatalism above.

Federman also rejects the *predictive* implication of determinism—such that *if* someone had full knowledge of all laws and conditions operant at any instant, they could *in principle* perfectly predict every event—on the ground that "all-knowing minds do not exist" (14). Disagreement about the Buddha's omniscience aside, whether anyone *attains* predictive omniscience is irrelevant. The predictive implication may be demonstrated by switching from the *epistemic* mode to the *entailment* mode: Determinism, if true, plus the conditions in the universe ages ago *entail* the choice one necessarily makes at a certain time *as a consequence* (regardless of whether anyone *knows* this particular entailment). This so-

[71] A "dharmic" action coheres with the Dharma; an "adharmic" action goes against it.

called "Consequence Argument" for *incompatibilism* is immune to epistemic limitations.[72]

Almost all of Federman's arguments about free will are from Dennett (*Varieties*), ignoring a variety of nuanced compatibilist positions that diverge from Dennett's and better cohere with Federman's account—particularly Frankfurt's ("Freedom"), whose influence on contemporary free will discussions is inestimable.[73] Frankfurt analyzes freedom *of the will* as having the sort of *will* one *wants*, an effective meta-will. Similarly, the Buddha prescribed the cultivation of a dharmic will through a variety of reflective means that yield an effective meta-will and constitute what he *did teach* as the path to *nirvāṇa*-F.[74]

Federman remarks that *nirvāṇa*, "the ultimate goal of the path, is referred to as freedom in the compatibilist sense" (9), but this is a hasty interpretation for three reasons: first, because *nirvāṇa*-F is obviously not the same kind of freedom discussed in connection with freedom of the will (where the question of compatibility pertains to determinism); second, because whether *nirvāṇa*-F is consistent with determinism is something that should not be assumed; and third, because it is a valid question whether *nirvāṇa*-F *conquers* or *transcends* determinism. Federman says Kalupahana describes *nirvāṇa* as "absence of constraint," but deflated-FW (as when a rat simply wants to run and does so) and *nirvāṇa*-F share absence of constraint. That deflated-FW and *nirvāṇa*-F share absence of constraint does not entail that *nirvāṇa*-F is compatibilist, for inflated-FW also shares absence of constraint, but is incompatibilist.

---

[72] The classical statement of this argument is in van Inwagen (*Essay*).

[73] See note 70 above.

[74] This is not to say that the Buddha advocated Frankfurt's type of model of freedom of the will, but rather that what he did advocate presupposes the ability Frankfurt identified and its importance on the Buddhist path.

Federman's reliance on the Pāli Canon is his prerogative, although it should be noted that from the Mahāyāna perspective he presents an incomplete account. Further, as noted in our discussion of Gier and Kjellberg, even some Pāli Buddhists had a more substantive view of agency than Federman's. Federman finally admits that the "Buddhist treatment of free will has to be extracted from the doctrine, as the doctrine is by no means a systematic philosophical treatise" (15), a keen disclaimer that would be placed more appropriately in his opening paragraph. In his ultimate paragraph he confidently asserts: "The kind of free will the Buddha taught is the acquired ability for clear reflection and wise choice" (15). But, again, although cyborg thermostats reflect on alternatives and make prudent choices, they seem clearly to lack free will in the moral-responsibility-entailing sense. Despite these limitations, Federman's analysis displays a wealth of canonical textual exegesis in support of the claim that the Buddha was *implicitly* compatibilist.

Harvey: More Wiggly (Pāli) Compatibilism

Peter Harvey's paper on the Theravāda Buddhist view of free will ("Freedom") makes arguments very similar to Federman's, minus Barnhart's fallacy, but adds some somewhat tangential but interesting canonical exegesis.[75] Thus, to avoid repetition, I only summarize Harvey's argument and call attention to certain issues. Harvey explicitly restricts his analysis to Theravāda (in this paper), and though he has doubts about the sort of free will that can obtain in the ultimately impersonal *skandhas*, he thinks that Theravāda supports compatibilism.

[75] This is not to suggest that Harvey makes no original contribution. He and Federman reference *each other's* papers.

Basically, Harvey reasons as follows. The Buddha's teachings rest on dependent origination, so it is plausible to think Theravāda Buddhism is determinist, though Harvey is uncertain whether determinism applies to enlightened beings (82-83). Harvey's explicit sensitivity to the elements of these issues that are uncertain is admirably Socratic. According to Harvey, the Buddha rejected fatalism because belief in *inevitabilism* implies volitional powerlessness over actions (paralysis of the will), and he advocated the importance of will in his prescriptions about dharmic action. Indeed, overall, he tends to see Theravāda Buddhism as more interested in whether people can break out of previous conditioned patterns in a self-directed way than in whether this breaking out is "free" in some philosophical sense.[76]

This emphasis is not unusual. Most Buddhists who think Buddhism has anything positive to say about anything related to free will tend to see the very notion of a Buddhist theory of free will as involving something like a category or syntax error, or some other kind of conceptual confusion.[77] Nonetheless, Harvey has a lot of positive things to say that would count toward a Buddhist theory of free will, were one to attempt its construction.

Like Federman, Harvey's account of the relevant elements of Buddhism that bear on the question of free will makes significant reference to the arguments of the Western analytic philosopher, Daniel Dennett, who is a compatibilist. Compatibilists in this debate, we may recall, are those who think it is logically possible both that determinism is true and that we sometimes exhibit free will. Some compatibilists (Dennett included) think that because agents may deliberate about undesired outcomes and choose actions that avoid them, the *lack of alternatives*

[76] For such an attempt, see Repetti ("Meditation").

[77] See, for example, Flanagan ("Quietism") and Garfield ("Mādhyamikas").

implied by determinism does not entail *inevitability*. The Buddha's teachings, on Harvey's interpretation, articulate the same volitional elements that do the explanatory work in Dennett's account, so we may conclude that the Buddha would have accepted wiggly-determinism.[78] As Harvey puts it, "the implied position of Theravāda Buddhism on the issue of 'freedom of the will' is a middle way between seeing a person's actions as completely rigidly determined, and seeing them as totally and unconditionally free" (86).

Echoing Siderits and Goodman,[79] Harvey also thinks the two truths and no-self doctrines imply there is no ultimate free will: "If there is no essential person-entity, 'it' can not be said to be either determined or free" (86). But he offers no explanation here, and his reasoning is analogous to the following (problematic, if not fallacious) claim: 'Shoes' cannot be said either to be made of leather or not made of leather because there are no ultimate shoes. For this critique to be fair, it would have to be formulated in terms of shoe *essences*, but I think the modified version of the objection, mutatis mutandis, would stand as well.[80] Needless to say, some shoes are not made of leather anyway.

Harvey's analysis is rich in canonical supports. Here is an illustrative gem Harvey has excavated (40):

> When someone said to the Buddha that there was no such thing as self-agency (*atta-kāra*), he replied by emphasizing that there is an "element of initiating (*ārabbha-dhātu*)" in

[78] I think that if he concluded with *soft* determinism, Harvey's argument would be stronger and more readily understood by Western philosophers.

[79] See my reviews of Siderits ("Reductionism") and Goodman ("Hard Determinism").

[80] See Repetti ("Reductionism"). Of course, Buddhist reductionism is as problematic as the free will issue, if not more problematic, so it is not a defect in Harvey's analysis that it makes reference to this doctrine.

> people—i.e., some kind of ability to choose—which allows them to initiate and direct actions such as bodily movements. . . . (AN III 337-338)

However, Harvey does not put this metaphorical diamond to use as a dialectical cutting tool. Though Federman seeks to weaken the implications of this passage about agency ("Buddha" n. 43), I consider it a diamond because it shows that the Buddha put the causal locus of action-origination in the individual's (non-inevitable) choice. Their apparent hesitance here is presumably because the very notion of "self-agency" seems to go against the Buddhist doctrine of the insubstantiality of the self, but if the agent/action pair is interdependently arisen (as per our review of Gier and Kjellberg's discussion of Nāgārjuna, above), no such threat arises.[81]

Harvey's tacit resistance to this line of thought might be due to the fact that the Buddha's reference to this "element of initiating" is illustrated by the possibly-*limited* ability to control one's movements: "For how can one who comes on his own (*sayaṃ abhikkamanto*) and returns on his own say 'There is no self-agency, there is no agency on the part of others'" (AN III 338). The ability to control one's movements is arguably what Frankfurt considers mere freedom of action. The question is whether the Buddha would extend this form of self-agency regarding one's actions to self-agency regarding the will—whether he would extend it, that is, from Frankfurt's freedom of action to Frankfurt's freedom of the will, the ability to have the sort of will one wants to have. I have argued that the Buddhist eightfold path outlines eight features of self-agency that enable the *ārya* to cultivate the sort of dharmic will she wants to have.[82] As long as the "self" in this Buddhist sort of "self-

[81] See Repetti ("Reductionism").

[82] Repetti ("Meditation").

agency" is understood in terms of psychological processes rather than as a metaphysically independent substance, there is no obvious inconsistency with Buddhist doctrine.

Some of Harvey's other remarks raise interpretive doubt. For instance, he says: "The Buddha opposed determinism as he saw it as a doctrine that froze a person's will to overcome unwholesome/unskillful (*akusala*) actions and develop wholesome/skillful (*kusala*) ones" ("Freedom" 40). Both he and Wallace quote Federman on this point, as if the Buddha discussed "determinism" in his objection to the *fatalist* view of Makkhali Gosāla (DN I 53-54). Goodman rightly objects to this inference, however, because the fatalism the Buddha rejected is not necessarily deterministic ("Resentment").[83] However, this passage arguably *does* support the Buddha's rejection of *hard* determinism, by implication, in that hard determinism entails an interpretation of *inevitability* that undermines the sort of proximal agency the Buddha seemed to be discussing in rejecting fatalism and that is sufficient for moral responsibility. This passage *does not* support the rejection of the *soft* determinist's interpretation of *evitability*, that allows that even though agents cannot alter whatever is determined to happen, the sort of proximal agential control individuals do exert over their own volitions, deliberative processes, choices, and actions is sufficient for purposes of moral responsibility. All recent-period scholars reviewed here quote Dennett, but they all also miss the elementary distinction that defines the debate: *hard* versus *soft* determinism, despite the fact that Dennett is a soft determinist. This is analogous to Western philosophers missing the elementary distinction between Theravāda and Mahāyāna Buddhism.

---

[83] Goodman is correct, but thinks Buddhism is hard determinist, and he ignores the fact that hard determinism and fatalism nonetheless share inevitabilism (one causal, one acausal, respectively); see Repetti ("Hard Determinism").

Harvey quotes Federman in support of what appears to be what I have described as a form of naïve autonomism, but one that is informed significantly neither by Buddhism nor by Western philosophical skepticism. It's as if, for them, "Buddhist free will" just means something weaker and more inchoate than deflated-FW, like "Buddhist who makes choices." Let's call this "naïve-FW," for even rats seem to make choices. (Recall the discussion of the "element of initiating" above, which Harvey suggests might be restricted to bodily movements.) And although he appeals to Federman's and Dennett's accounts for *some* support, what is needed is an account of why we should think mere *choice-making* differentiates between hard and soft determinism and supports the latter: hard determinists do not deny we make choices, but that we are free to choose *otherwise* under identical deterministic conditions.[84] Keeping with the metaphor, despite the many dialectical gems Harvey uncovers, he does not sufficiently construct explanatory jewelry.

The remainder of Harvey's work revolves around the following panoply of intuitively relevant issues: how Buddhism regards madness and social and biological conditioning; the relationships between karma, the self, dependent-origination and freedom; and the relationships between willing and spiritual ignorance, attention, "the basic radiant purity of the mind," and spiritual freedom. Although much of this exhibits a wealth of meritorious canonical-gem-excavation, most of these points Harvey explores do not straightforwardly support his argument; however, they may be understood to function in a way similar to what

[84] The soft determinist thinks we *would be* free to choose otherwise under slightly different deterministic conditions: *if we wanted to* choose otherwise; the hard determinist replies we *can never want to* choose otherwise; the soft determinist counters we *could* want to choose otherwise *if* we had different wants; the hard determinist retort is that we can *never* have different wants; and so on. The debate boils down to whether it is reasonable to consider unactualized possibilities sufficient for purposes of moral responsibility. I argue that it is reasonable (*Counterfactual*).

we saw with Gier and Kjellberg's identification of terrain in need of further charting by interpretive cartographers. Harvey devotes significant attention (47-53) to the questions whether only the Buddha or other enlightened beings could be vulnerable to the effects of past karma, whether everything that dependent origination covers is karmic (alternately put, whether karma and dependent origination are coextensive),[85] and whether the earliest canonical sources or slightly-later *Abhidharma* agree on these points. He follows this section with significant attention (54-61) to whether karma is fatalistic (it seems fairly obviously not fatalistic, but it is to his credit that he gives the canonical supports for this view), stating that "karmic results of a particular action are actually seen to vary, so past karma does not inflexibly determine a fixed result, produced in a mechanical-like way" (59)—more wiggly-determinism.

Regarding the issue of the self, Harvey reviews the arguments of Siderits, Goodman, and others, and (correctly, I think) complains that there is a two truths equivocation, but he does not specifically identify it. Harvey seems to think the individual is not a "myth" *at the conventional level* (62) and that its agency is efficacious (autonomous):

> As the mind is ever-changing, subject to a variety of internal and external conditions, it is seen as good to develop greater guidance over the way it operates, based on wise restraint and an understanding of how it operates. Thus the wayward mind can be controlled, by understanding dawning within it, so that, so to speak, it takes more responsibility for itself and the actions it brings about, such that these are more coherent with the

[85] As I argued above, despite some disagreement on this, the Buddha's restriction of the karmic to the volitional seems to count strongly against the equation of the karmic with the causal.

> genuine benefit of the overall pattern of mental states. (67)

These remarks may be used to support a model of meta-mental volitional regulation and, to the extent that—both within and outside Buddhism—a thing's causal powers count as grounds for its ontological status, they also support a somewhat Pudgalavādin-type view that the Buddhist has intermediate-level functional and/or ontological status (between conventional/ultimate) as a self-regulating (autonomous) agent. Harvey disputes this reading, arguing that all that he discusses in the quotation above may be accounted for as the working of processes within the *skandhas*.[86] I am not claiming that his remarks *imply* my reading, but only that they *may be used* to support it. Setting aside the question about the possibility of affording middle-level ontological status to the self-agency-exhibiting psychological processes Harvey's remarks might be used to support, let us focus only on the ability, and call it "Buddhist-FW." Buddhist-FW shares elements of deflated-FW and naïve-FW and is consistent with my earlier soft determinist trifurcation: virtually-libertarian, virtually-hard, and relatively-soft. Let me review this distinction.

What I describe as *virtual libertarianism* is only virtually or functionally equivalent to libertarianism, not actually equivalent, because it is still deterministic. Libertarianism is the view that we possess the sort of full-fledged freedom of the will most ordinary folks think they possess when they think their choices originate entirely autonomously and are ultimately fully up to them. But this ability is only virtual if we live in a deterministic world (because for it to be truly ultimately up to us, we would need to live in an indeterministic world in which our choices were not determined long before we were born). Enlightened beings are virtu-

[86] Harvey (communication, April 2013).

ally libertarian because, lacking any self-sense or ego-volitions that could be thwarted, nothing that could undermine their free will determines their behavior.

What I describe as *virtual hard determinism* differs from actual hard determinism, which asserts that because determinism is true, everyone lacks the sort of agency that would count as free will or ground moral responsibility. Virtual hard determinism is only virtual for worldlings who are so exogenously determined by conditions of ignorance and ego-volitions that it is *as if* they lack any endogenous agency altogether, even the most minimalistic type.

What I describe as *relative soft determinism* differs from soft determinism *simpliciter*, which asserts only that determinism is true and agents sometimes exhibit free will. Relative soft determinism differs from this simply in terms of acknowledging a full spectrum of graded degrees to which agents are exogenously and/or endogenously determined in ways that are conditioned more or less heavily by ignorance and ego-volitions. This latter category captures the bulk of *āryas* as well as others who through various efforts have cultivated some sort of agential self-regulative ability, in varying degrees.

All three categories are forms of soft determinism. For, they all allow for the truth of determinism and its compatibility with varying degrees of free will, even that of virtual hard determinism, which is not *actually* hard determinism because *virtually* hard determined agents nonetheless retain the potential for liberating behavioral changes, say, upon hearing the Dharma.

In concluding my review of Harvey's account, I must admit that two points of Harvey's that I have underplayed are the importance he

attaches to the liberating role of attention,[87] and that no one thing is ever "the cause" of anything because things arise from particular clusters of conditions. Related to his emphasis on attention is the important fact that mental phenomena are not just what they are, say, like physical things are, but that they are also *intentional*—that is, they have objects that they are *about*; their nature opens out onto other things, and *how* one sees things is a potent mental condition for self-changes. I think these ideas are rich in theoretical potential, but we cannot develop them here, other than to suggest that they intuitively support Buddhist-FW as well as my own version of a Buddhist theory of free will.[88]

Lastly, it should be noted that although it is not emphasized in his paper, as regards spiritual freedom Harvey seems to think this is made possible by someone, say, A, becoming spiritually free, teaching B, and B intently listening and acting on this (75-76). This helps others break out of restricting patterns of conditions. This path of breaking out is still a collection of conditioned and conditioning processes, yet what is realized at the culmination of this path, *nirvāṇa*, is unconditioned. This makes it possible to construe the sort of agency the Buddhist path tacitly

[87] See Joerg Tuske ("No-Self") for an attempt to locate an Indian philosophical basis for free will in an analysis of attention. Tuske criticizes an element of my similar view as set forth in Repetti ("Meditation")—not the part of my analysis that emphasizes the centrality of attention, but that part that borrows distinctions from Frankfurt. His objection is a standard objection against Frankfurt's model, but because I do not use the model in the way Frankfurt did, Tuske's objection does not really undermine my account.

[88] Repetti ("Meditation") argues basically that a plausible Buddhist conception of free will may be construed as a function of the sort of meta-cognitive and meta-volitional attention-training practices that constitute the Eightfold Buddhist Path.

endorses minimally, even deterministically, if need be, as simply a function of hearing the Dharma.[89]

Wallace: Mahāyāna Incompatibilist Free Will?

B. Alan Wallace, in "A Buddhist View of Free Will: Beyond Determinism and Indeterminism," describes enlightened behavior as exhibiting a "kind of freedom" such that "one nonconceptually rests in this timeless, pristine awareness, allowing actions to arise spontaneously and effortlessly, aroused by the interplay of one's own wisdom and the needs of sentient beings from moment to moment" (67-68). And whereas Federman and Harvey emphasize the conventional validity of free will, Wallace flatly rejects the *compatibilist's* deflated-FW conception on the grounds of what he takes to be the obvious truth of the sort of incompatibility between determinism and free will expressed in the Consequence Argument. (Briefly, the Consequence Argument holds that if determinism is true, then all choices are unalterable, necessary "consequences" of prior, nomologically sufficient events functioning as causal conditions, in which case there cannot be free will; that is, determinism and free will are incompatible, so, if determinism is true, it must be hard determinism.) For more subtle reasons he also seems to reject the *incompatibilist's* inflated-FW, but he apparently accepts other forms of free will, as shall be explained shortly.

Wallace makes many interesting claims that may be adopted by an autonomist, about how Buddhist practices like mindfulness meditation decrease bondage, moving the practitioner closer to *nirvāṇa*-F, as I have also argued ("Meditation"). Because they are similar to those Har-

[89] Harvey (communication, April 2013).

vey and Federman made to advocate the opposite (compatibilist) view, I will ignore them here in order to highlight what is distinctive in Wallace's views.

Wallace treats determinism as if it must be hard *and* fatalistic: "Fatalism is the unavoidable implication of determinism as surely as later events are inevitably set in stone by prior conditions" ("View" 65). Because the Buddha rejected fatalism and Wallace sees determinism as implying fatalism, he rejects determinism, despite Goodman's objection against equating them, a considerable literature on *soft* (non-fatalistic) determinism, and Federman's counter-argument. (Wallace references Federman, so he is presumably aware of these objections.) But he also rejects *indeterminism*, because both opposing doctrines presuppose a kind of metaphysical realism rejected in Mahāyāna (65).

The anti-realist (if not idealist) Mahāyāna idea that Wallace seems to have in mind, simplifying greatly, is that there is no conceptual-construction-independent or independently existing substantive reality, but both deterministic and indeterministic models depict reality as a conceptual-construction-independent, independently existing substantive or objective reality: on the deterministic view, a conception-independent reality is constituted by a sequence of discrete events all of which are causally necessitated; on the indeterministic view, a conception-independent reality is constituted by a sequence of discrete events many of which are not causally necessitated. Another important element of Wallace's model is the idea that consciousness is inherently unobstructed, free, boundless, and so on. So, if the Mahāyāna model is correct, as Wallace thinks it is, then that would explain why he would say that *if* determinism was true, *then* free will is impossible: because *if* determinism was true, *then* consciousness would not seem to be unobstructed, free, boundless, and so on, either.

What kind of clear understanding can we expect to attain, however, if both determinism and its opposite are equally non-indicative of reality—if reality escapes logical bivalence (the theory that every meaningful proposition is either true or false, but not both)? In developing this view, Wallace makes insightful parallels between an allegedly non-linear holographic model of interdependent origination, and paradoxical conceptions of time and the like, suggesting a multiply-complex ultimate (*non-conventional*) metaphysics shared by Mahāyāna and physics that renders the standard approach to free will otiose. But, as noted above, interdependence—holographic or otherwise—arguably implies mega-linear-determinism. The evidence Wallace offers for these models, if any, is insufficient to displace determinism, indeterminism, or both.

Let us try to unpack some complex Mahāyāna Buddhist ideas Wallace relies on in his attempt to explain his own conception of free will. Although several different terms are employed in that literature, such as "buddha-nature," "brightly shining mind," "mind of clear light," "primordial consciousness," and so on, these are all relatively synonymous. Basically, the idea suggested by these terms is that we are already what enlightened beings realize that we are, although we are obstructed from that realization due to our afflictions. The Buddhist path removes the obstructions. Although we may not realize it, we experience our buddha-nature at times when the afflictions are thoroughly disabled, such as between thoughts, in dreamless sleep, at death, and so forth. It is not obvious from his presentation, but Wallace's argument may be understood as the claim that all sentient beings have, at this presently inaccessible level, complete freedom, understood to include or involve something analogous to if not identical with free will. He seems to argue further that this freedom can be experienced in tantric practice, where, for example, the mind realizing emptiness appears in form as a deity

(Buddha) and when one has divine pride (complete identification with the deity). One then is manifesting enlightened consciousness.[90]

Wallace refers to the *brightly shining substrate consciousness* (*ālaya-vijñāna*), a way in the Tibetan Buddhist tradition of referring to buddha-nature, present at all times, even between attention-engaged moments, in dreamless sleep, at death, and in the *bardo* (post-mortem) state (for enlightened beings), noting "This realm of consciousness is beyond the scope of the conceptual mind, so its possible influence on the minds of ordinary sentient beings is unimaginable" (63).[91] Wallace claims that the substrate consciousness is the same non-dual trans-temporal consciousness that "lies beyond the realm of philosophy" (63), shared by the Buddha and one's own "future" enlightened self (i.e., that one already has it, but it is rarely manifest). In the practice of *divine pride* (the essence of tantric deity yoga, the strong identification with the deity), "one draws the transformative power of one's future enlightenment into the present moment, with the understanding that the future is not inherently real and separate from the present" (67). So far, this does not suggest reverse temporal causation, but perhaps "positive thinking."

Nevertheless, Wallace seems to bite the reverse-causation bullet: "In such practice, based on a realization of emptiness and the Buddha nature of all beings, one is free to enable the future to influence the present" (67). However, to interpret Wallace as accepting reverse tem-

---

[90] I owe some of the interpretive suggestions in this paragraph to Daniel Cozort (communication, December 2013). See Rinpoche (*Chariots*) for an in-depth explanation of these tantric practices and ideas.

[91] It must be noted that this claim bears a resemblance—hopefully, only a superficial one—to the Biblical theodicy to the effect that because God's ways are beyond human comprehension, it is appropriate to suspend doubt in the face of difficulties in comprehending God's intentions relative to the state of affairs God permits in the world. One important difference is that the theodicy seems ad hoc.

poral causation here would overstate his understanding. Arguably, the idea is that the ability to imaginatively be a Buddha in the future is what can influence the present, and thus can enlighten. However, there is much in the literature Wallace cites and in his many other writings that suggests a conception of time and the external world as mind-dependent in a way that might ground what may be described as titanic psychic abilities. It is not clear, therefore, whether Wallace means to interpret divine pride and related notions in a naturalistic or in a more titanic, supernatural sense. If he intends the latter, however, such a titanic reverse-temporal power *far exceeds* the non-physical-self model that grounds his own rejection of inflated-FW (59-60). Let's call this robust tantric sort of free will "titanic-FW."

If Wallace only associated these titanic-FW powers with enlightened beings, this might not be as implausible. But Wallace claims that the Great Perfection (Dzogchen, Tibetan Buddhist) teaching posits the substrate consciousness as what not only enables "a kind of freedom that transcends the demarcations of past, present, and future," but as what undergirds reincarnation (68)—something to which all unenlightened beings are subject. This is also problematic, if not contradictory, because reincarnation is *by definition* temporal, and the Buddhist view of the worldling is that he is stuck in the *kālachakra* (temporal wheel, of rebirth). Further, if substrate consciousness enables reincarnation and reverse-causation, and is ever-present between moments of attention-absorption and during dreamless sleep, then everyone shares elements of titanic-FW (while deeply asleep or attention-disengaged)—something much more inflated than the inflated-FW Wallace rejects.

Naturalistic western libertarians would reject titanic-FW, ironically, as unrealistic. I am not suggesting that these tantric ideas—fantastic as they might appear from a Western perspective—are false, but only pointing out an inconsistency. In fairness, Wallace could

respond that although ordinary folks are thought to possess buddha-nature, and even to experience it (without realizing it as such) in certain moments, such as in between thoughts or in dreamless sleep, they are not thought to possess *siddhis* (psychic powers), even if *āryas* and advanced masters are thought to attain them under certain circumstances. That counter seems correct, but I doubt it resolves the inconsistencies noted above. It leaves open the possibility that *āryas* attain a state in which they are exempt from physical laws.

Wallace proposes "a modern definition of freedom as the capacity to achieve what is of value in a range of circumstances" as what worldlings have in mind as free will (62), and claims they lack it. (Surely, however, *āryas*—titanic or not—do not lack it). His "modern definition," however, resembles the sort of freedom of action or *naïve-FW* discussed above in connection with the accounts of Federman and of Harvey. Wallace seems to contrast *nirvāṇa*-F with naïve-FW when he says "the Buddhist tradition clearly emphasizes that ordinary sentient beings are *not* free, for we are constrained by mental afflictions such as craving, hostility, and delusion" (62). Wallace seems to be saying that we lack naïve-FW and *nirvāṇa*-F, and only advanced Buddhists have a substantive free will: "freedom of will depends on the ability to recognize the various impulses that arise involuntarily in the mind and to choose which among them to accept or reject" (64). This resembles Buddhist-FW. It seems for Wallace that worldlings possess elements of titanic-FW, but not inflated-FW, deflated-FW, naïve-FW, Buddhist-FW, or *nirvāṇa*-F;[92]

[92] Recall that titanic-FW is the sort possessed by almost mythically supernatural-powers-possessing *āryas* and *Arhats*, illustrated by divine pride's trans-temporal causation. Inflated-FW is the sort libertarians think we possess, such that our choices are entirely autonomous and ultimately up to us, thought to be impossible in a deterministic world and often associated with a non-physical conception of the self or soul. Deflated-FW is the sort of minimalist view that Harvey and others hold, such that the *skandhas* exhibit certain self-regulative abilities. Naïve-FW is the pre-philosophical view

āryas possess titanic-FW and Buddhist-FW; and enlightened beings possess *nirvāṇa*-F (and elements of titanic-FW). But do enlightened beings possess the sort of individual-agent-type volitional regulation found in Buddhist-FW? Wallace's analysis skirts around this *agentless-autonomy* issue, but this is perhaps *the* central Buddhist version of the free will problem—how a non-agent can be autonomous in any meaningful sense.

Frankfurt's meta-volitional regulation account sounds strikingly "Buddhist" in light of Buddhist-FW and Wallace's related claims that "metacognitive awareness allows for the possibility of freely choosing whether or not to allow a desire to lead to an intention or to let an intention result in verbal or physical action," and that mindfulness "entails directing one's attention to wholesome and unwholesome tendencies and recognizing them as such so that one may cultivate the former and reject the latter" (64). These descriptions also mirror the Theravādin criteria, but Frankfurt's and the Theravādin's accounts are compatibilist. Thus, because the principle of parsimony (or lightness) favors the simpler, less metaphysically risky hypothesis, there seems to be no reason why these compatibilist abilities require any type of incompatibilist indeterminism or any determinism/indeterminism dichotomy-circumscribing, bivalence- and conceptuality-transcending stance, apart from the fact that the latter describes Wallace's overall metaphysical stance independently of this particular issue, as may be seen in most of his other works.[93]

---

that a being has free will if it makes choices and/or controls its bodily movements in action. Buddhist-FW is the sort of free will *āryas* possess in varying degrees as they increasingly cultivate the ability to regulate their volitions in accordance with the Dharma. And *nirvāṇa*-F is the liberation achieved upon attainment of enlightenment.

[93] See, for example, Wallace (*Dimensions*; *Training*; and *Revolution*).

Wallace rejects compatibilism on the ground that determinism is fatalistic (59, 61-62, 65), but, as I argued above, *soft* determinists need more than the allegation of fatalism to displace their dialectical credibility as evitablists, and—although this has apparently escaped his attention—Wallace's own proposal of Buddhist-FW is *prima facie* determinism-friendly. Wallace rejects indeterminist (libertarian) inflated-FW on the ground that the Buddha rejected pure chance (59-60), but Robert Kane and other libertarians have sophisticated versions of indeterminism that differentiate it from "pure chance."[94]

Wallace rejects metaphysical realism on the grounds of quantum holographic interdependence and related elements of Mahāyāna metaphysics (65-66), but, as I argued above, interdependence arguably implies mega-linear-determinism, and trans-temporality and other titanic powers implicitly require indeterminism *within* the realm of conditioned phenomena. He adds the conceptuality-transcending substrate consciousness (67) in place of what he rejects, but because *all of these ideas require greater support than determinism or indeterminism*, his account begs the question—at least from a non-Mahāyāna or a non-Buddhist vantage.

All things considered, and perhaps to his credit, Wallace's concluding position is so complex—it interweaves affirmations and

[94] Kane ("Pathways"). "Soft indeterminists" are indeterminists who reject the argument that indeterminism is "hard" (incompatible with free will) on the ground that randomness that undermines agency. They argue that indeterminism could play a useful role, for example, in the generation of a greater number of considerations of alternate possibilities that might become cognitively available to a deliberator than might be available under a deterministic version of that process, but so long as the deliberator selects from among them on the basis of his or her considered values and reasons, the mere presence of the element of chance somewhere in the deliberative process does not entail that the decision is purely random.

denials of all the forms of freedom discussed, relativized to worldlings, Buddhists, and enlightened beings—as to defy subsumption within any singular standard Western taxonomy, such as "libertarian" or "incompatibilist." But it seems fair to describe most of it as consistent with my trifurcated compatibilist model, because agentless virtual-libertarianism fits enlightened beings who exhibit *nirvāṇa*-F; relative-soft determinism fits *āryas* who exhibit Buddhist-FW; and everyone else is virtually-hard-determined, because (contra Wallace) they possess *too much* of a form of naïve-FW that may be described as *unrestricted volitional expression* (which fosters mental bondage) and thus they significantly lack deflated-FW, Buddhist-FW, and *nirvāṇa*-F, and any titanic-FW they might technically satisfy in substrate consciousness is thoroughly unconscious, not voluntarily accessible, and thus presumably causally and/or functionally impotent.

Wallace's essay contains many rich and potentially fruitful ideas, but—not unlike the other recent-period scholars reviewed here—he fails to sufficiently scaffold them to support the sort of philosophical weight they are supposed to bear. Despite their interesting character, many of his claims are in greater need of support than, say, plausible versions of *soft* determinism (such as deflated-FW or Buddhist-FW) or even certain plausible forms of libertarianism (inflated-FW), such as Kane's. I suggest that the missing philosophical supports may be found scattered throughout Wallace's many other exceptional works on the convergence between Buddhism and science, in seed form, waiting to be watered and cultivated. That would certainly be a worthwhile project. It is one that I hope is undertaken in the near future, whether by Wallace or another Mahāyāna philosopher.

Conclusion

Recall that this article is the fourth in a four-part series, and that this conclusion will contain reference to the entire series. Thus, review of early-, middle-, and recent-period scholarship reveals that with regard to free will, Buddhist thought is as complex as Western thought. Some of the complexity is the way Western scholars of Buddhism try to tease out implications of passages in Buddhist texts; some is the way Buddhist scholars try to extract elements of thought consistent with Buddhist conceptions from Western philosophical texts. Because Buddhism has only fairly recently begun to articulate conceptions of free will, more sophisticated positions may be expected to emerge. It is a very interesting and pregnant phase of scholarship,[95] so it would be misleading, therefore, to seek "the" Buddhist theory of free will, just as it would be to seek "the" Western theory of free will.

One conclusion shared by most of the extant Buddhist scholarship, except that of Goodman, is that Buddhism presupposes some sort of free will for the Buddhist aspirant, if not for the worldling. Another shared conclusion, excepting Goodman and Wallace, is that because the Buddha rejected inevitabilism, he would reject the view that determinism is hard (incompatible). That is, because the Buddhist path presupposes volitional dynamics that soft determinists deploy, if the Buddha would have equated determinism with dependent origination, he would have accepted soft determinism; if not, he would have accepted a close second, what may be called "soft dependent origination," the view that dependent origination is compatible with deflated-FW or Buddhist-FW.

---

[95] I am currently working on an edited collection of approximately twenty (mostly) original articles on Buddhism and free will, tentatively titled "*Agentless Agency? Classical and Contemporary Buddhist Perspectives on Free Will*," as well as a monograph, tentatively titled "*Buddhism and Free Will*," both currently under review at Routledge.

Mahāyāna and Theravāda divergence did not play a significant role in early- and middle-period accounts, but recent-period scholarship *mostly* supports separate treatments. I say "mostly" because my own recent-period theory is an exception to this doctrinal division.[96] Although elements of my own account have appeared throughout my treatment of these three periods of scholarship, I have not directly reviewed it here; but because my view is compatibilist and attempts to be pan-Buddhist syncretic, I see it as consistent with both Theravāda and Mahāyāna. I argue, in short, that Buddhism presupposes free will for the same reasons Harvey and Federman do, but I draw considerable support for this claim from an analysis, largely informed by Frankfurt's meta-volitional model,[97] of the causal dynamics of cultivating attention to volition presupposed by the Buddhist meditative path.

In summary, Mahāyānists' non-realist metaphysics circumvents the determinism/indeterminism dichotomy, but begs the question in the free will debate (insofar as it presents a more complex puzzle as a "solution" to a simpler one), and *interdependent* origination does not circumvent linear determinism, but suggests *mega-linear*-determinism. Theravāda Buddhists express what I have described as a form of *wiggly*-determinism, an attempt to circumvent the implications of hard determinism that ignores the distinction between hard and soft determinism, but on my analysis dependent origination is perfectly consistent with *soft* determinism, even if the two doctrines are not otherwise identical. The insubstantial-self and reductionist doctrines are more explicit for Mahāyānists and understood more radically by them, but that understanding has not been sufficiently set forth or supported

[96] See Repetti ("Meditation").

[97] The Frankfurt-informed causal analysis of free will that is applied to the Buddhist view is the subject of Repetti ("Meditation"), but it is worked out in great, mostly non-Buddhist, detail in Repetti (*Counterfactual*).

in terms of its implications for the free will question, although no argument presented here suggests that it cannot be set forth.

Some of the remaining unresolved central issues are whether dependent origination is determinism, whether it applies to enlightened beings, whether the insubstantial-self doctrine implies there is no autonomy because there is no agent or allows agentless autonomy, and how reductionism applies to free will. Although Buddhism has only begun to articulate its many possible stances on free will, many potentially explanatory dialectical gems have been uncovered. Indeed, Buddhism has only begun to reveal a glimpse of this diverse potential.[98]

Bibliography

Adam, Martin T. "No Self, No Free Will, No Problem: Implications of the *Anattalakkhaṇa Sutta* for a Perennial Philosophical Issue." *Journal of the International Association of Buddhist Studies* 33, no. 1 (2011), 239-265.

Aronson, Harvey B. *Buddhist Practice on Western Ground Reconciling Eastern Ideals and Western Psychology.* Boston, MA: Shambhala Publications, 2004.

Conze, Edward. *Buddhist Thought in India.* London: George Allen and Unwin, 1962.

Davidson, Donald. *Essays on Actions and Events.* Oxford: Clarendon Press, 1980.

[98] For various attempts to address additional, selective elements of these issues, see Adam ("No Self"), Flanagan ("Quietism"), Frequinon ("Buddhism"), Garfield ("Mādhyamikas"), McCrae ("Emotions"), Meyers (Freedom), Repetti (Panel; "*Dharma*-responsive"), and Siderits (Panel).

DeGrazie, David. *Animal Rights: A Very Short Introduction.* New York: Oxford University Press, 2002.

Dennett, Daniel. *Elbow Room: The Varieties of Free Will Worth Wanting.* Cambridge, MA: MIT Press, 1984.

Federman, Asaf. "What Kind of Free Will Did the Buddha Teach?" *Philosophy East and West* 60, no. 1 (January 2010), 1-19.

Flanagan, Owen. "Buddhist Quietism about Free Will." Presentation at the Columbia Society for Comparative Philosophy (November 9, 2012). Online. Available HTTP: < http://www.cbs.columbia.edu/files/nov-9-flanagan.m4a> (accessed July 20, 2013).

Frankfurt, Harry. "Freedom of the Will and the Concept of the Person," *Journal of Philosophy* 68 (1971), 5-20.

_____. "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66, no. 23 (1969), 829-839.

Frequinon, Marie. "Buddhism and Free Will." Presentation at the Columbia Society for Comparative Philosophy (April 24, 2009). Online. Available HTTP: <http://www.cbs.columbia.edu/24-APR-FRIQUEGNON-%28Trans-cut%29.pdf> (accessed July 20, 2013).

Garfield, Jay. "Just Another Word for Nothing Left to Lose: Freedom, Agency and Ethics for Mādhyamikas." In *Freedom of the Will in a Cross-Cultural Perspective*, edited by E. Bryant. New York: Oxford University Press, forthcoming. Online. Available HTTP: <http://www.smith.edu/philosophy/documents/NothingLefttoLose.pdf> (accessed July 17, 2013).

_____. "Nāgārjuna's Theory of Causality: Implications Sacred and Profane," *Philosophy East & West* 51, no. 4 (October 2001), 507-524.

Gier, Nicholas F. and Paul Kjellberg. "Buddhism and the Freedom of the Will: Pali and Mahayanist Responses." In *Freedom and Determinism*, edited by Joseph Keim Campbell, Michael O'Rourke, and David Shier, 277-304. Cambridge, MA: MIT Press, 2004.

Goodman, Charles. "Resentment and Reality: Buddhism on Moral Responsibility," *American Philosophical Quarterly* 39, no. 4 (October 2002), 359-372.

Gowans, Christopher W. *Philosophy of the Buddha. New York: Routledge*, 2003.

Harvey, Peter. *An Introduction to Buddhism: Teachings, History and Practices*, 2d ed. Cambridge, MA: Cambridge University Press, 2012.

_____. "'Freedom of the Will' in the Light of Theravāda Buddhist Teachings," *Journal of Buddhist Ethics* 14 (2007), 35-98.

_____. *The Selfless Mind: Personality, Consciousness and Nirvana in Early Buddhism.* Richmond, UK: Curzon Press, 1995.

Hume, David. *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton. Oxford, U.K.: Clarendon Press, 2007.

Kalupahana, David J. *Buddhist Philosophy: A Historical Analysis*. Honolulu: University Press of Hawaii, 1976.

Kane, Robert. "Some Neglected Pathways in the Free Will Labyrinth." In *The Oxford Handbook of Free Will*, edited by Robert Kane, 406-437. New York: Oxford University Press, 2002.

Lehrer, Keith. "Freedom and the Power of Preference." In *Freedom and Determinism*, edited by Joseph Keim Campbell, Michael O'Rourke, and David Shier, 47-69. Cambridge, MA: MIT Press, 2004.

McCrae, Emily. "Emotions, Ethics, and Choice: Lessons from Tsongkhapa," *Journal of Buddhist Ethics* 19 (2012), 344-369.

Meyers, Karin L. Freedom and Self-control: Free Will in South Asian Buddhism. Dissertation: University of Chicago, December 2010.

Nozick, Robert. *Philosophical Explanations.* Cambridge, MA: Harvard University Press, 1981.

Priestly, Leonard. "Pudgalavāda Buddhist Philosophy," *The Internet Encyclopedia of Philosophy*. Online. Available HTTP: <http://www.iep.utm.edu/p/pudgalav.htm> (accessed July 17, 2013).

Rick Repetti. "The *Dharma*-responsive Theory of Free Will: Enough 'Freedom to' to Attain 'Freedom from.'" Presentation. Columbia Society for Comparative Philosophy. December 7, 2012. Online. Available HTTP: <http://www.cbs.columbia.edu/cscp/2012/11/dec-7-rick-repetti-the-dharma-responsive-theory-of-free-will-enough-freedom-to-to-attain-freedom-fro.html> (accessed July 17, 2013).

_____. "Buddhist Hard Determinism: No Self, No Free Will, No Responsibility," *Journal of Buddhist Ethics* 19 (2012), 130-197.

_____. "Buddhist Reductionism and Free Will: Paleo-compatibilism," *Journal of Buddhist Ethics* 19 (2012), 33-95.

_____. Panel. "Contemporary Perspectives on Buddhist Ethics" conference, Columbia University (October 7, 2011). Online. Available HTTP: <http://www.cbs.columbia.edu/buddhist_ethics/panel-five.html> (accessed July 17, 2013).

_____. "Earlier Buddhist Theories of Free Will: Compatibilism," *Journal of Buddhist Ethics* 17 (2010), 279-310.

_____. "Meditation and Mental Freedom: A Buddhist Theory of Free Will," *Journal of Buddhist Ethics* 17 (2010), 24-69.

_____. *The Counterfactual Theory of Free Will: A Genuinely Deterministic Form of Soft Determinism.* Saarbrücken, Germany: LAP Lambert Academic Publishing, 2010.

Rinpoche, Khenchen Palden Sherab. *Turning the Wisdom Wheel of the Nine Golden Chariots*. Sidney Center, New York: Dharma Sumudra, 2009.

Siderits, Mark. Panel. "Contemporary Perspectives on Buddhist Ethics" conference, Columbia University (October 7, 2011). Online. Available HTTP: <http://www.cbs.columbia.edu/buddhist_ethics/panel-five.html> (accessed July 17, 2013).

_____. "Paleo-Compatibilism and Buddhist Reductionism," *Sophia* 47, no. 1 (2008), 29-42.

_____. *Personal Identity and Buddhist Philosophy: Empty Persons.* Burlington, VT: Ashgate Publishing, 2003.

_____. "Beyond Compatibilism: A Buddhist Approach to Freedom and Determinism," *American Philosophical Quarterly* 24, no. 2 (April 1987), 149-159.

Stcherbatsky, T. *The Central Conception of Buddhism*. New Delhi: Motilal Barnarsidass, 1974.

_____. *Buddhist Logic.* The Hague: Mouton and Company, 1958.

Thakchoe, Sonam. "The Theory of Two Truths in Tibet," *The Stanford Encyclopedia of Philosophy* (Summer 2011). Online. Available HTTP: <http://plato.stanford.edu/archives/sum2011/entries/twotruths-tibet/> (accessed January 25, 2014).

Tuske, Joerg. "The No-Self Theory and Problems in the Philosophy of Mind." In *A Companion to Buddhist Philosophy*, edited by Steven M. Emmanuel, 419-428. Malden, MA: Wiley-Blackwell, 2013.

Van Inwagen, Peter. *An Essay on Free Will.* Oxford: Clarendon Press, 1983.

Wallace, B. Alan. "A Buddhist View of Free Will: Beyond Determinism and Indeterminism." In *Neurosciences and Free Will*, edited by Robert Pollack, 59-69. New York: The Earth Institute at Columbia University, 2008.

_____. *Hidden Dimensions: The Unification of Physics and Consciousness.* New York: Columbia University Press, 2007.

_____. *The Attention Revolution: Unlocking the Power of the Focused Mind.* Somerville, MA: Wisdom Publications, 2006.

_____. *The Seven-Point Mind Training.* Ithaca, NY: Snow Lion Publications, 2004.

Woodward, James and Christopher Hitchcock. "Explanatory Generalizations, Part I: A Counterfactual Account," *Noûs* 37:1 (2003), 1-24.